Software Requirements: Update, Upgrade, Redesign

Hoorn, J.F.

2006

document version
Publisher's PDF, also known as Version of record

Software Requirements: Update, Upgrade, Redesign.
Towards a Theory of Requirements Change

SIKS Dissertation Series No. 2006-14
The research reported in this thesis has been carried out under the auspices of SIKS, the Dutch Research School for Information and Knowledge Systems.

Dit onderzoek is verricht met een subsidie van Senter*Novem*, het agentschap van Economische Zaken, in het kader van de Innovatieve Onderzoeksprojecten (IOP), projecttitel *Integrating Design of Business Processes and Task Analysis,* projectnummer MMI99009.

Promotiecommissie:
prof.dr. Hans van Vliet (promotor)
prof.dr. Gerrit van der Veer (promotor)
prof.dr. Ian Sommerville (St. Andrews University, UK)
prof.dr. Per Runeson (Lund University, Sweden)
prof.dr. Roel Wieringa (Universiteit Twente)
prof.dr. Sjaak Brinkkemper (Universiteit Utrecht)
dr. Jaap Gordijn (Vrije Universiteit Amsterdam)

CIP-GEGEVENS KONINKLIJKE BIBLIOTHEEK, DEN HAAG

Hoorn, J.F.

Software Requirements: Update, Upgrade, Redesign.
Towards a Theory of Requirements Change
Thesis Vrije Universiteit
-With ref.- With summary in Dutch.

ISBN 90-9020925-5

Subject headings: Software, Requirements, Business, Computer-Human Interaction

VRIJE UNIVERSITEIT

Software Requirements: Update, Upgrade, Redesign

Towards a Theory of Requirements Change

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor aan
de Vrije Universiteit Amsterdam,
op gezag van de rector magnificus
prof.dr. L.M. Bouter,
in het openbaar te verdedigen
ten overstaan van de promotiecommissie
van de faculteit der Exacte Wetenschappen
op maandag 9 oktober 2006 om 13.45 uur
in het auditorium van de universiteit,
De Boelelaan 1105

door

Johannes Ferdinand Hoorn

geboren te Haarlem

promotoren: prof.dr. J.C. van Vliet
prof.dr. G.C. van der Veer

Acknowledgments

My research of the past four years concerned change. Changing is all I did during my career. I worked in the faculty of Literature and Arts, Psychology, Science, Earth- and Life Sciences, and now I am teaming up with Ordina, a business company for ICT, consultancy, and outsourcing. I changed my research topics quite a few times, from metaphor in poetry to fictional characters to software requirements to e-learning in science communication. I changed quite a bit in research methods, from hermeneutics to reaction time and EEG recordings, from neural net analysis to ethnography and questionnaire design. Yet, none of this would have been possible without the help of a considerable number of people. For my time at Computer Science, the following people supported, facilitated, judged, published or downright participated in conducting the research (in alphabetical order).

Michiel Bakker, ConQuaestor
Jelle van den Berg, MSc Computer Science, VU
Mark Breuker, MSc Computer Science, VU
Arnoud M. de Bruin, Policy Consultant Medical Center
Jos Burkx, Business Unit Manager, Ordina System Integration & Development
Cristina Chisăliță, PhD Computer Science, VU
Simon van Dam, AHB Ltd (former Vice President of VirTouch Ltd)
Anton Eliëns, Intelligent Multimedia, VU
Tom M. van Engers, Professor Legislative Knowledge Management, UvA
Guido Fambach, The Mediator Group
Jo Geraedts, Océ Industrial Design Department
Rik Gerrits, LibRT B.V.
Johan de Heer, Telematica Institute
Zhisheng Huang, Artificial Intelligence, VU
Kees Jong, MSc Computer Science, VU
Geurt Jongbloed, Stochastics, VU
Martijn Kamermans, President ConQuaestor
Bob Klaasse, MSc Computer Science, VU
Evelien Kok, Concern Information Management Police
Ger Koole, Professor Business Mathematics and Computer Science, VU
Hans Kruithof, Senter*Novem*
Elly Lammers, Office Manager, Computer Science, VU
Brenda Neuteboom, ICT IDM Management
Robin Nieuwenhoff (paranymph), Nieuw Financieel Advies
Arco van Nieuwland, Cofounder of Exact Software
Sandra Pronk, MSc Business Mathematics and Computer Science, VU

Mari-Carmen Puerta Melguizo, Content and Knowledge Engineering, UU
Joost Schalken, PhD Software Engineering, VU
Radboud A. W. Vanlerberghe, O&I Management Partners
Peter Veerman, Head Anesthesiology
Pieter-Paul Verheggen, Motivaction
Gerrit van der Veer, Professor Interaction Design, VU (the man who brought me in)
Hans van Vliet, Professor Software Engineering, VU
Henriette van Vugt (paranymph), PhD Computer Science, VU
Marc Woolderink, Business Unit Manager, Ordina System Integration & Development
Jan van Wouwe, Computer Systems, VU
and numerous participants in the field experiments.

I also want to thank the following people for their support during the last year (in alphabetical order):

Mieke Beer, Project Manager, Science Communication, VU
Mark Bos, PhD Science Communication, VU
Joske Bunders, Professor Biology and Society, VU
Cees Koolstra, Communication Science, VU
Roy Kloet, MSc Science Communication, VU
Antoinette Liemburg, PhD Science Communication, VU
Jolanda Veldhuis, PhD Science Communication, VU
Jaap Willems, Professor Science Communication, VU (the man who saved me)

But above all, I thank Elly Konijn, my indestructible partner in research and life.

The change that impacted my life the most, however, was the birth of our twins Lieke and Nina. They taught me how to bend the system – my own, that is.

Contents

Executive Summary		3
Samenvatting		6
1 Introduction		11
1.1	What's the Problem?	11
1.2	Requirements Change	13
1.3	Reasons for Requirements Change	14
1.3.1	Understanding the Problem	14
1.3.2	Communicating the Problem	14
1.3.3	Outside Events	14
1.4	Scope and Rationale	15
1.4.1	Scope	15
1.4.2	Rationale	16
1.5	Overview of the Book	17
References		18
2 Theory		23
2.1	General Background: DUTCH Design	23
2.2	Groupware Task Analysis	25
2.3	Business Situations	27
2.4	Integrating Business Models with Task Analysis	28
2.5	The 'Engineering' of Requirements	30
2.6	Task-based Worlds – Task-based Emotions	32
2.7	Goal Types: Personal vs. Business, Egotistic vs. Altruistic	33
2.8	Aspects of Processing: Speed-accuracy Trade-off	34
2.9	Outline: Model of Requirements Change (MoRC)	35
2.10	Adaptation: Change of Stakeholder Requirements Model	38
2.11	Research in Businesses	40
2.11.1	Different Business Models	40
2.11.2	Problems of Controlled Field Work	42
2.11.2.1	Science Critique	43
2.11.2.2	Business Critique	44

2.11.3	Conclusions/Discussion	45
2.12	The Remainder of this Book	47
References		47
3 Relevance and Prioritization		51
3.1	Introduction	51
3.2	Study 1: Goal Relevance and Agreement to Requirements	53
3.2.1	Method	54
3.2.1.1	Participants and Experimental Design	54
3.2.1.2	System	55
3.2.1.3	Procedure	56
3.2.1.4	Measurements	57
3.2.2	Analysis and Results	58
3.2.2.1	Scale Analysis	58
3.2.2.2	Agreement to Goal Relevance and Requirements	59
3.2.3	Discussion of the Focus Switch	62
3.3	Study 2: Changes in Requirements Prioritization	63
3.3.1	Method	65
3.3.1.1	Participants and Experimental Design	65
3.3.1.2	Procedure	67
3.3.1.3	Measurements	67
3.3.2	Analysis and Results	69
3.3.3	Discussion of the Prioritization Experiment	70
3.4	Study 3: Changes in Requirements Prioritization Revisited	71
3.4.1	Method	72
3.4.1.1	Participants and Experimental Design	72
3.4.1.2	System	74
3.4.1.3	Procedure	74
3.4.1.4	Measurements	74
3.4.2	Analysis and Results	75
3.4.2.1	Prioritization Analysis	75
3.4.2.2	Relevance of Goals	77
3.4.3	Discussion of Prioritization, Old and New Goals, and Conservatism	78
3.5	General Conclusions and Discussion	79
3.5.1	It’s Okay for the Business but not Necessarily for Me	79
3.5.2	Message to Practitioners	82
References		83

Appendix 3.1 86
Appendix 3.2 118
Appendix 3.3 121

4 Valence and Agreement 123

4.1 Introduction 123
4.2 Theory 125
4.2.1 The Type of Goals 125
4.2.2 Valence 125
4.2.3 Not Only Must Haves 126
4.2.4 Variability in Agreement 126

4.3 Study 1: Viewpoints on Requirements and Goals Unaligned 127
4.3.1 Method 127
4.3.1.1 Participants and Experimental Design 127
4.3.1.2 System 128
4.3.1.3 Procedure 128
4.3.1.4 Measurements 129
4.3.2 Analysis and Results 132
4.3.2.1 Scale Analysis 133
4.3.2.2 Agreement to CMS Requirements and Goals 133
4.3.2.3 Discussion of the Effects on Agreement in the CMS Case 136
4.3.2.4 CMS: Management-viewpoint Requirements in Relation to Personal Goals 139
4.3.2.5 Discussion of the Relation between Management-viewpoint Requirements and Personal Goals 140
4.3.2.6 CMS: Management-viewpoint Requirements in Relation to Business Goals 141
4.3.2.7 Discussion of the Relation between Management-viewpoint Requirements and Business Goals 143
4.3.3 CMS: Practitioner's Perspective 145

4.4 Study 2: Business Viewpoints on Requirements and Goals Aligned 146
4.4.1 Method 147
4.4.1.1 Participants and Experimental Design 147
4.4.1.2 System 147
4.4.1.3 Procedure 147
4.4.1.4 Measurements 149
4.4.2 Analysis and Results 149
4.4.2.1 Scale Analysis 149

4.4.2.2	Agreement to LWMS Requirements and Goals	150
4.4.2.3	Discussion of the Effects on Agreement in the LWMS Case	151
4.4.2.4	LWMS: Management-viewpoint Requirements in Relation to Business Goals	152
4.4.2.5	Discussion of the Relation between Management-viewpoint Requirements and Business Goals	153
4.4.3	LWMS: Practitioner's Perspective	154
4.5	Study 3: Personal Viewpoints on Requirements and Goals Aligned	156
4.5.1	Method	157
4.5.1.1	Participants and Experimental Design	157
4.5.1.2	Procedure	157
4.5.1.3	Measurements	159
4.5.2	Analysis and Results	159
4.5.2.1	Scale Analysis	159
4.5.2.2	Agreement to COTS PCs Requirements and Goals	160
4.5.2.3	Discussion of the Effects on Agreement in the COTS PCs Case	161
4.5.2.4	COTS PCs: Personal-viewpoint Requirements in Relation to Personal Goals	161
4.5.2.5	Discussion of the Relation between Personal-viewpoint Requirements and Personal Goals	163
4.6	Study 4: Personal Viewpoints on Requirements and Goals Revisited	165
4.6.1	Method	165
4.6.1.1	Participants and Experimental Design	165
4.6.1.2	System	165
4.6.1.3	Procedure	166
4.6.1.4	Measurements	169
4.6.2	Analysis and Results	169
4.6.2.1	Scale Analysis	169
4.6.2.2	Agreement to VTPlayer Requirements and Goals	169
4.6.2.3	Discussion of the Effects on Agreement in the VTPlayer Case	170
4.6.2.4	VTPlayer: Personal-viewpoint Requirements in Relation to Personal Goals Revisited	171
4.6.2.5	Discussion of the Relation between Personal-viewpoint Requirements and Personal Goals Revisited	172
4.6.3	VTPlayer: Practitioner's Perspective	173
4.7	General Conclusions and Discussion	174
4.7.1	The Goals-to-Requirements Chiasm	175

4.7.2	The Requirements-analysis Rift	177
4.7.3	The Status of Valence	177
4.7.4	Methodological Challenges	178
4.7.5	Message to and Testimonies from Practitioners	179
References		181
Appendix 4.1		184
Appendix 4.2		190
Appendix 4.3		194
5 Usability and Efficiency		197
5.1	Introduction	197
5.2	Usability as a Concept	198
5.3	Usability and its Relation to Satisfaction	200
5.4	Usability and its Relation to Performance	201
5.5	Stakeholder Logistics	202
5.6	The Human and the Machine Side of Performance	203
5.6.1	Effectiveness (output)	204
5.6.2	Efficiency (throughput)	206
5.6.3	Effort (input)	210
5.7	Relevance and Valence	213
5.8	Conclusions on Theory Development	214
5.9	Study 1: Effectiveness, Efficiency, Usability, and Satisfaction with 25 Banking Systems	215
5.9.1	The Usability-Satisfaction Hypotheses	215
5.9.2	Method	215
5.9.2.1	Participants, Systems, and Procedure	215
5.9.2.2	Measurements	216
5.9.3	Analysis and Results	216
5.9.3.1	Testing the ISO 9241-11 Definition of Usability	216
5.9.3.2	Testing Stakeholder Logistics	218
5.9.4	Conclusions for the ISO 9241-11 Definition of Usability versus Stakeholder Logistics	219
5.9.5	The Efficiency Hypothesis	220
5.9.6	Method	220
5.9.6.1	Measurements	220
5.9.7	Analysis and Results	220
5.9.7.1	Testing the Efficiency Hypothesis	220
5.9.7.2	Testing the Speed-Accuracy Hypothesis	220
5.9.8	Conclusions Efficiency Hypothesis	221
5.9.9	Discussion Study 1: Interactive Banking Systems	222
5.10	Study 2: Efficiency, Usability, and Satisfaction in the Operating Room	223

5.10.1	Method	224
5.10.1.1	Participants	224
5.10.1.2	Business Process and Systems	225
5.10.1.3	Procedure	227
5.10.1.4	Measurements	227
5.10.2	Analysis and Results	229
5.10.2.1	Scale Analysis	229
5.10.2.2	Hypothesis Testing	231
5.10.3	Conclusions/Discussion Study 2: Operation Room Systems	236
5.11	General Discussion	239
References		241
Appendix 5.1		245
Appendix 5.2		252

6 Conclusions 255

6.1	Back to CoStaR	255
6.1.1	Variability in Agreement	256
6.1.2	The Goals-to-Requirements Chiasm	257
6.1.3	The Type of Goals	257
6.1.4	Relevance-Valence	258
6.1.5	The Requirements-analysis Rift	258
6.2	Revised CoStaR	259
6.3	Back to DUTCH	262
References		264

Executive Summary

This investigation is about a spoiled child that made a wish list for her birthday but once she got the presents, longed for something different or at least in another color and wrapped in different paper. We all are this spoiled child when we try to build something or have it built for us. Making a cabinet in that difficult corner of the living takes several rounds of negotiation and redesign. The builder really gets annoyed when the cabinet is half finished and hears that it should be done all over again, because “our new plasma wide-screen takes much more room than the old telly.”

When something like this happens while a large information system is built - and it does, a lot - redesigning the system becomes more costly, frustrating, and time-consuming the later in the process the request for a change occurs. The literature states that anticipating requirements change in early design, knowing where to look for it, which requirements are more vulnerable than other, is worthwhile to save costs, spare on frustration, and reduce overall development time. There are several approaches to this problem, packed under the name of requirements engineering and in particular requirements change. These approaches are used to detect inconsistency between requirements, to discover tacit requirements, or resolve conflicting requirements, which can all give rise to requirements change. However, the common requirements engineering approaches usually focus on just one specific system. To describe and analyze requirements, these approaches often apply formal logics that - according to certain authors - are not intuitively clear to the future users or other stakeholders of the system. Moreover, validation of the requirements by means of hard-nosed empirical testing and heavy weight statistics is rare within the requirements engineering domain.

The present study tackles requirements change with empirical tests and statistic verification of general hypotheses that apply to more than just one specific system. In an attempt to seek out the mechanisms that underlie requirements change, the main tool used was a range of structured questionnaires that systematically queried the relations between requirements and the types of stakeholder goals that gave rise to those requirements. Asking a question is intuitively clear to people but the way in which these questions were structured made it possible to translate the answers into a scientific model.

To see whether the general hypotheses I formulated really applied to all kinds of systems, in all kinds of organizations with all kinds of users and other stakeholders, I investigated a Capacity Management System at the Dutch police, the Web-based e-learning system Didactor at the Science Faculty and at ConQuaestor financial consultancy, a Logistic Warehouse Management System at a provincial government in The Netherlands, Commercial Off-the-Shelf

PCs with expert users, the VTPlayer Braille mouse at Bartimeus College and VISIO with blind children, 25 banking systems at a multinational bank in The Netherlands, and operation room equipment (i.e. a new input module at the anesthesiologist's device) with anesthetists in an academic medical center.

I based my research on a few assumptions. If everybody agrees that a car should have wheels, it will have wheels. If everybody disagrees that a car should have wooden tires, the car won't have wooden tires. However, if a car should run on alcohol to avoid pollution, the opinions are divided. Some will say that alcohol is easy to produce but raises tax problems; others may say that pollution is not right but that harming the oil industry is worse. Thus, what people argue about most (here, alcohol as car fuel) is most vulnerable to change. That the car must have wheels and won't have wooden tires are stable requirements, alcohol is a volatile requirement.

One of the major empirical findings in my studies with software requirements was that people argued most about what stuff to get rid off in order to achieve something. An example is to get rid of Windows (supposedly a user-friendly shell but error prone) in order to improve the system's stability. Similarly, people also argued a lot about what things were needed to keep them out of trouble. For example, managers who wanted to install anti virus software in printers that were not connected to the Internet (and thus, could not be infected). Hence, *the requirements most vulnerable to change were positive requirements that were connected to negative goal states and negative requirements connected to positive goal states.* This peculiar mechanism I coined the 'goals-to-requirements chiasm.' It occurred independent of type of system, type of organization (government, financial, medical, educational), or type of stakeholder (novices, experts, students, managers, workforce). Thus, knowing in advance which must requirements are raised to avoid negative situations and knowing which won't requirements are raised to achieve desirable situations is important to detect the areas of volatile requirements - chances are high that these will become subject to a change request.

However, as a second important empirical finding, no one will ever find the volatile requirements at an early stage if goals and requirements are not explicitly connected. This seems like an obvious stand but all too often, stakeholders are confronted with a list of loose requirements with the question whether this is what they want or not. However, during my research I found that stakeholders - even the expert ones - hardly oversee the consequences of a proposed system feature for their future work situation. So most of the time they will state that they more-or-less agree with a requirement - the safe choice. Yet, once you spell out what the pros and cons are for achieving goals with the system or not, so I noticed during my investigations, stakeholders are far more outspoken in their agreement or disagreement with a requirement and whether to put it high or low on the priority list. Because analysts often forget to relate requirements to goals (a traceability problem), I coined this situation the 'requirements-analysis rift.' The danger of not bridging this rift is that the

requirements change occurs late in the project or worse, after the system is implemented. Only then stakeholders experience what the system actually does or does not do for their work. Early prototyping may be a solution - although academically, I think that is a bit ad hoc.

Nonetheless, prototyping does not help either if the prototype simulates a system that merely supports business goals. I repeatedly established the effect that the way the system affects personal goals is far more important than how it affects business goals - and I guess that this is even true for the business managers themselves. One of the results was that personal goals were regarded as more relevant, and management-viewpoint requirements were disagreed to more from a personal than a business perspective. More importantly, stakeholders showed significantly more changes in the prioritization of requirements if they prioritized from a personal viewpoint. Thus, a software-development project is bound to fail if its orientation is on the business goals alone.

Knowing, then, that requirements should be connected to and analyzed for their effects on particularly personal goals, that requirements attached to goals of opposite polarity are most sensitive to change, I wondered which type of goal would be the ultimate source for a change request. Being brought up in a tradition of user-centered design (the DUTCH approach), my angle was that usability problems would lead to dissatisfaction of the stakeholder and that hence, requirements change transpired. This was not so. It turned out that for professional use, stakeholders were most concerned with the efficiency of their machines, in particular the fact that the machine should not make errors. Estimates of machine efficiency considerably explained why stakeholders thought their machines were usable and why they were satisfied with their machines or not. As another manifestation of the goals-to-requirements chiasm, inaccuracy of the machines as a goal to avoid best predicted the level of agreement to the must requirements. Human accuracy as a goal to approach best predicted the won’t requirements. In other words, Windows may be user-friendly in certain respects but because it is error prone, it is still regarded as not-so-usable and things MUST CHANGE. The challenge for us user-centered designers then is to accommodate a bunch of system-centered users.

This investigation is about stakeholders of an interactive system that make wish lists for their future IT but once the development starts, long for systems that make fewer errors and that serve their personal goals better. If the must haves change, the stakeholder probably is driven by fear. If the won’t haves change, the stakeholder probably is driven by desire. However, if the wish list does not explain what the wishes actually are for, the child cries once she gets the goods and starts playing with them.

Samenvatting

Dit proefschrift is getiteld *Eisen aan Computerprogrammatuur: Bijwerken, Opwaarderen, Herontwerpen. Naar een Theorie over Veranderingen in het Pakket van Eisen.* Maar eigenlijk gaat dit onderzoek over een verwend kind dat een verlanglijstje maakte voor haar verjaardag, maar toen ze eenmaal de pakjes kreeg, verlangde naar iets anders of in ieder geval in een andere kleur en ingepakt in een ander papiertje. Wij allen zijn dit verwende kind wanneer we iets willen maken of het voor ons laten doen door een ander. Een kast maken in die lastige hoek daar in de woonkamer vereist de nodige onderhandelingen en pogingen tot herontwerp. De kastenbouwer raakt geïrriteerd als de kast al half af is en dan hoort dat alles overnieuw moet omdat “ons nieuwe plasma breedbeeldscherm veel meer ruimte inneemt dan onze oude TV.”

Als iets dergelijks gebeurt tijdens de ontwikkeling van een groot informatiesysteem - en dat gebeurt, vaak zelfs - wordt het herontwerpen van het systeem steeds kostbaarder, frustrerender en tijdrovender naarmate een verzoek tot verandering later in het traject gedaan wordt. Het kunnen anticiperen op verandering van eisen in een vroeg stadium van het systeemontwerp, weten waar je moet zoeken, welke eisen gevoeliger zijn voor verandering dan andere, is belangrijk om kosten te besparen, ter voorkoming van frustratie en om de algehele ontwikkeltijd te bekorten. Er zijn verschillende benaderingen van dit probleem, die bekend staan onder de naam van “vereistenconstructie” en in het bijzonder “vereistenverandering.” Hoe belangrijk deze benaderingen ook zijn om inconsistenties tussen eisen op te sporen, verborgen eisen te expliciteren, of conflicterende eisen te herkennen - die immers allemaal kunnen leiden tot veranderingen in het eisenpakket - ze richten zich slechts op één specifiek systeem and maken vaak gebruik van formele logica’s die niet intuïtief duidelijk zijn voor de toekomstige gebruikers en andere belanghebbenden in het systeem. Bovendien is validatie van het eisenpakket door middel van harde empirische toetsing en zware statistische tests zeldzaam binnen het domein.

Het voorliggende proefschrift bestudeert veranderingen in het pakket van eisen met behulp van proefnemingen en statistische tests. Het toetst algemene hypothesen die verder reiken dan slechts één bepaald systeem. Om de mechanismen op te sporen die ten grondslag liggen aan veranderingseisen was mijn voornaamste onderzoeksinstrument een reeks van gestructureerde vragenlijsten die systematisch de verbanden tussen vereisten en doelen van de belanghebbenden in kaart brachten. Een vraag stellen is intuïtief te begrijpen voor mensen maar de manier waarop de vragen waren gestructureerd maakte het mogelijk de antwoorden te vertalen in een wetenschappelijk model.

Om te zien of de algemene hypothesen die ik had opgesteld ook werkelijk golden voor allerlei systemen, voor verschillende organisaties met alle moge-

lijke gebruikers en andere belanghebbenden, onderzocht ik een Capaciteitsbeheersysteem bij de politie, het Web-gebaseerd electronisch leersysteem Didactor bij de Faculteit der Exacte Wetenschappen en bij ConQuaestor financieel adviseurs, een Logistiek Opslagbeheersysteem bij een provinciale instelling, PCs die zo van de plank samengesteld kunnen worden bij expert-gebruikers, de VTPlayer Braillemuis bij het Bartimeus College en VISIO Amsterdam met blinde kinderen, 25 bankierssystemen in een multinationale bank uit Nederland en apparaten uit de operatiekamers (met name het anesthesie-apparaat) in een academisch medisch centrum.

Als iedereen het erover eens is dat een auto wielen moet hebben, dan zal de auto wielen hebben. Als iedereen het erover oneens is dat een auto houten banden moet hebben, dan zal de auto geen houten banden hebben. Echter, als een auto op alcohol moet lopen om luchtvervuiling te voorkomen, raken de meningen verdeeld. Sommigen zullen zeggen dat alcohol eenvoudig te maken is maar wel tot belastingproblemen leidt; anderen zullen zeggen dat luchtvervuiling niet goed is maar dat het schaden van de olie-industrie nog erger is. Waar de mensen dus het meest over debatteren (in dit geval over alcohol als brandstof) is het meest onderhevig aan verandering. Dat een auto wielen moet hebben en geen houten banden zijn stabiele vereisten. Dat de auto op alcohol moet lopen is een instabiele vereiste.

Een van de belangrijkste vondsten bij de eisen die men aan een softwarepakket stelt, was dat mensen de meeste onenigheid hadden over dingen die ze kwijt wilden om een bepaald doel te bereiken. Een voorbeeld is dat Windows (gebruiksvriendelijk maar foutgevoelig) moest verdwijnen om de stabiliteit van het systeem te verbeteren. Mensen hadden ook onenigheid over dingen die er nodig bij moesten om moeilijkheden te voorkomen. Managers wilden bijvoorbeeld antivirusprogramma's voor printers die niet met het Internet verbonden waren (en dus ook helemaal niet geïnfecteerd konden raken). Dus, *de eisen die het gevoeligst voor verandering zijn waren positieve vereisten die verbonden waren met (het vermijden van) negatieve situaties en negatieve vereisten die verbonden waren met (het tot stand brengen van) positieve situaties.* Dit merkwaardige mechanisme noemde ik het doelen-en-vereisten chiasme (kruisstelling, overkruising). Dit chiasme trad op onafhankelijk van het type systeem, type organisatie (overheid, financieel, medisch, educatief), of type belanghebbende (nieuwelingen, experts, studenten, leidinggevenden, gewone medewerkers). Om veranderingsgevoelige vereisten op te sporen is het dus belangrijk om van te voren te weten of wat er in het systeem moet voortkomt uit de angst voor negatieve situaties en of wat er uit het systeem weg moet voortkomt uit het verlangen naar een positieve situatie – de kans is groot dat juist deze eisen zullen veranderen.

Echter niemand zal instabiele eisen vroegtijdig kunnen vinden als doelen en eisen niet nadrukkelijk met elkaar verbonden zijn. Dat lijkt voor de hand liggend maar al te vaak krijgen belanghebbenden een lijst van losse eisen voor hun neus met de vraag of dit nou is wat ze willen of niet. Maar zelfs de experts

kunnen nauwelijks overzien wat een voorstel voor een (eigenschap van het) systeem in de praktijk voor hun werk betekent. Dus meestentijds zal men aangeven het er wel zo'n beetje mee eens te zijn – de veiligste keuze. Als je echter voorkauwt wat de voors en tegens zullen zijn voor een toekomstige werksituatie, zijn de belanghebbenden ook meteen veel uitgesprokener in hun mening over een systeemeigenschap en geven ze duidelijker aan of het hoog op de prioriteitenlijst moet of niet. Omdat analisten vaak vergeten om eisen aan doelen te koppelen (een traceerbaarheidsprobleem), heb ik dit verschijnsel de vereisten-analyse breuk genoemd. Het gevaar van het laten bestaan van deze 'breuk' is dat de veranderingen in het eisenpakket pas laat in het ontwikkeltraject naar boven komen of erger, nadat het systeem al geïmplementeerd is. Pas dan ervaren de mensen wat het nieuwe systeem eigenlijk voor hun werk betekent. Prototypes kunnen ten dele het probleem verhelpen, maar meestal wordt een prototype vrij laat in het ontwikkelproces gebouwd en academisch gezien is het een beetje een lapmiddel achteraf. Je zou veranderingseisen willen kunnen voorspellen.

Bovendien, prototypes helpen niet als het prototype een systeem weerspiegelt dat alleen bedrijfsdoelen ondersteunt. Ik heb in mijn proeven herhaaldelijk vastgesteld dat de wijze waarop het toekomstig systeem de persoonlijke doelen beïnvloedt veel belangrijker is dan hoe het de bedrijfsdoelen beïnvloedt – en ik denk zelfs dat dat ook geldt voor de persoonlijke doelen van de directeuren van het bedrijf. Belanghebbenden in het systeem vertoonden significant meer veranderingen in de prioritering van vereisten als ze dat deden vanuit een persoonlijk standpunt. We kunnen dus stellen dat een ontwikkeltraject gedoemd is te mislukken als het systeem zich slechts richt op het halen van bedrijfsdoelen.

Wetende dat systeemeigenschappen gekoppeld moeten worden aan en geanalyseerd moeten worden voor hun effecten op met name persoonlijke doelen en dat vereisten gekoppeld aan doelen met tegengestelde polariteit (positief-negatief) het meest gevoelig zijn voor verandering, vroeg ik me af welk soort doel nou de sterkste bron van veranderingseisen zou zijn. Omdat mijn achtergrond in het mens-georiënteerde ontwerpen ligt (de DUTCH benadering) was mijn insteek dat bruikbaarheidsproblemen leiden tot ontevredenheid bij de belanghebbenden en dat om die reden veranderingseisen zouden worden gesteld. Maar dit was niet het geval. Het bleek dat in de beroepspraktijk, mensen vooral gericht waren op de efficiëntie van hun machines en het meest beducht waren voor systeemfouten. Schattingen van machine-efficiëntie verklaarden tot op grote hoogte waarom mensen dachten dat hun machines bruikbaar waren en waarom ze tevreden waren met hun machines of niet. Opnieuw deed het doelen-en-vereisten chiasme zich voor. Inaccuratesse van de machines als situatie die vermeden moest worden voorspelde het beste de mate waarin men instemde met de eigenschappen die het systeem moest hebben. Menselijke accuratesse, een wenselijke situatie, voorspelde het beste de eigenschappen die het systeem niet moest hebben. Met andere woorden, Windows mag dan in

sommige opzichten gebruiksvriendelijk zijn, maar omdat het foutgevoelig is, wordt Windows nog steeds gezien als niet-zo-bruikbaar en MOETEN er dingen veranderen. De uitdaging voor ons mens-geörienteerde ontwerpers is dus om tegemoet te komen aan een stelletje systeemgerichte gebruikers.

Dit onderzoek gaat over belanghebbenden van een interactief systeem die een wensenlijst maken voor hun toekomstige IT maar die, wanneer eenmaal het ontwikkeltraject in gang is gezet, verlangen naar systemen die minder fouten maken en beter afgestemd zijn op hun persoonlijke doelen. Als de dingen die erin moeten veranderen, is de belanghebbende waarschijnlijk gedreven door angst. Als de dingen die eruit moeten veranderen, is de belanghebbende waarschijnlijk gedreven door verlangen. Echter, als de wensenlijst niet aangeeft waar de wensen eigenlijk voor zijn, zal het kind huilen als ze eindelijk haar speelgoed krijgt en ermee gaat spelen.

1 Introduction

"It is not the strongest of the species that survives, nor the most intelligent, but rather the one most responsive to change." Attributed to Charles Darwin (Wikiquote, 2005)

Abstract

The main concern of this book is to answer the question why requirements on a computer system change. Implementing changes is time-consuming and cost-ineffective. This chapter gives an outline of the problematic, summarizes a number of answers that can be derived from the literature, and previews how the research presented in subsequent chapters is conducted. In short, stakeholders change their minds about what they want from a system because a market event has happened, because their understanding of the problem that the system should help solve evolved, or because the communication of the problem was imperfect. In this book, I attempt to offer new insights in why and how requirements change. I discovered two phenomena that can be active during requirements evaluation and that impact the extent to which stakeholders agree upon a set of requirements. The goals-to-requirements chiasm (Chapter 4) explains that stakeholders argue about the desirable features of a future IT most when these features are related to goals to avoid (sic) with the system. As a complement, stakeholders argue most about unwanted aspects of a future IT when these are related to the goals to achieve with the system (sic). In professional use, the main concern in this respect is to keep the systems error free (Chapter 5). Communication breakdowns can occur when the requirements-analysis rift (Chapter 3) is operative. Although on the one hand, stakeholders may disagree to a list of requirements from a viewpoint of personal work experience, they hold their tongue when that list is presented to accomplish the necessary business goals. In a series of field experiments, I attempted to fuse proper quantitative research with real business cases. In Chapter 2, I defend that a scientific and empirical approach to software and requirements engineering is useful to establish best practices that are reliable as well as generally applicable.

Keywords: Requirements change, Stakeholders, Business

1.1 What's the Problem?

It is hardly an overstatement that in the industrialized countries, a vast number of interactive systems have entered everyday life and that they are not about to

leave soon. Today, we make a call on a mobile phone, listen to an MP3 player, retrieve money from an automated teller machine, or fill out an electronic tax form. Users sometimes hardly realize that they interact with a computer, let alone that they know how many development cycles, versions, and re-releases have preceded that ordinary tool they have become so accustomed to.

System developers do. They struggle with evolving markets, technological breakthroughs, new legislation, and the latest fashion hype. A major problem in developing a system is to know what functionality a system should offer, what user goals it should support or what business processes it should facilitate. Requirements engineering (RE) is a series of organized activities to obtain and document such knowledge for system engineers as well as for other stakeholders who are involved in developing or using the system (e.g., the client, managers, end-users, and maintenance personnel). Yet, the motto of the IEEE RE'05 Conference was this citation of Brooks (1987):

> The hardest single part of building a software system is deciding precisely what to build. No other part of the conceptual work is as difficult as establishing the detailed technical requirements, including all the interfaces to people, to machines, and to other software systems. No other part of the work so cripples the resulting system if done wrong. No other part is more difficult to rectify later. (Brooks, 1987)

The problem gets worse when stakeholders change their minds about what they want from the system. Particularly when a system is under development, a change request can have serious impact on the design of a system (cf. Coakes & Elliman, 1999). The business situation sometimes changes so quickly that change requests repeatedly occur during the course of development. According to Boehm (1981), a regular software development track for large information systems shows about 25% of changes. Redesign, however, is expensive, time-consuming, and often frustrating. Boehm and Pappacio (1988) calculated that changes in the later stages of software development of large systems cost 50 to 200 times more than the same changes made in the early design stages. After the system is operational, the costs for implementing a change request increase to a 1,000 times the costs in early design. For smaller systems, the damage is a little less severe (Boehm & Basili, 2001). It is nevertheless important that we can anticipate requirements change. If we know why requirements change and what mechanisms govern change, it might be possible to detect 'the danger zones' – the requirements most susceptible to change – in the early stages of requirements elicitation and gathering.

However, we are dealing with rapid changes. Therefore, we not only need to know which requirements on a specific system in a specific business case are changing and why, we also need generic knowledge on requirements change. With this type of knowledge, we – hopefully – can anticipate changes while being less dependent of the particular system under construction and less vulnerable to the time aspect.

1.2 Requirements Change

Why do you want something different than before? “Stable requirements are the holy grail of software development” (McConnell, 1993) but requirements change, whether we like it or not. The reality being modeled in a tax-handling system changes when tax laws change. The information-needs of the user of an automated library system change when new types of knowledge carriers emerge. The management information that a bank manager needs, changes after a take-over or major reorganization. Users change or evolve their needs when they have experienced the possibilities and impossibilities of the systems they work with.

Van Vliet (2000, p. 214) indicates that many software developers and requirements engineers work from a 19th century, Taylorian perspective of ‘scientific management.’ This approach has shown valuable in the rationalization of industrial processes. Tasks are broken down into smaller tasks and through tests and observations, the manager can find the one best solution to accomplish that task. This one best way is then standardized in the form of rules and procedures. Requirements engineers often work in the same way. They gather information about (computer) tasks, acquire requirements through observation of the work floor, and interview domain experts as well as other stakeholders. The assumption is that there is a static list of requirements, which ‘only needs to be found’ during the analysis and does not need to be checked with the stakeholders (ibid.). In this line of thought, requirements can be ‘engineered’ just like software components can.

We all know this to be a fallacy (Hsia et al., 1993). Over the years, a number of improvements over the traditional ‘just tell me what you want’ approach has been proposed. Ethnography (Simonsen & Kensing, 1997; Viller & Sommerville, 2000) and other approaches from the social sciences have entered the arena. Complementary, socially oriented approaches to system development have been proposed, in which the analyst operates as a change agent (Beyer, 1995). In these approaches, reality is not something ‘out there,’ but is constructed during the process (Checkland, 1999). More recently, agile and other lightweight processes have been proposed to deal more effectively with the changing world around us (Kovitz, 2003). All of these developments in many ways improve the requirements engineering process and its outcome. Yet, requirements change. They will change after the system has become operational. Quite likely, they will also change during the development process.

However, where do these changes come from? At the surface level, requirements change seems whimsical. We might even claim that requirements engineering has much in common with weather forecasting in that in both cases there is a limit as to how far the future can be predicted.

1.3 Reasons for Requirements Change

Some of the reasons why requirements change have to do with understanding the problem that should be solved by technology, communicating it between developers and stakeholders, and events that happen outside the scope of the technological system. I will give a quick overview.

1.3.1 Understanding the Problem

In many cases, the problem that is 'modeled' by the new system is not clear-cut (cf. Lowe, 2003). The problem space or task domain that is modeled by an automated teller machine is more deterministic than the problem space of a group decision support system. Obviously, people can learn (cf. Edberg & Olfman, 2001). Therefore, developers as well as other stakeholders can have a different view of the problem while their understanding evolves. In addition, stakeholders can hardly imagine their future work situation. They sometimes do not know themselves what they want from the system. As a result, communicating the problem becomes difficult and certain requirements are not captured in the early stages of design.

1.3.2 Communicating the Problem

During requirements elicitation and capturing, certain requirements may be left unnoticed. Sometimes, tacit knowledge or implicit suppositions are not made explicit because stakeholders find them so obvious. If the analyst does not pose the right questions in the beginning or forgets to ask for feedback from the work floor, these requirements may turn up at a later stage. Therefore, requirements documents should be frequently updated. However, even if updated regularly, engineers and developers can interpret the requirements differently from other stakeholders. This may be due to ambiguous language or a different understanding of the problem. One way to improve the communication is to have feedback and requirements validation repeatedly, which can be done by running a prototype or playing use cases (e.g., Leffingwell & Widrig, 2003). Another reason why requirements seem to change unexpectedly is because there is no control loop on the process of change itself (the need for change management) (Wiegers, 2003, pp. 280-296).

1.3.3 Outside Events

Edberg and Olfman (2001) reported that change requests on maintenance work at a sample of organizations were for 60% directed at the functionality (because of business model change, new technology, learning, and legislation). Only 15% of the change requests were directed at repair work such as debugging. External events can cause changes in the initial problem that was supposed to be solved by the new technology. Due to market events, politics, a merger, new technology or a new law, business models can change (e.g., a software production house becomes a service-oriented organization). New

goals are to be met with the technology or new business processes may need to be supported, which can call for a change. Another threat is that the project's budget runs dry or that the timeline is shortened (cf. WorkSmart™ Guide, 2005) so that the cheapest and quickest solutions are wanted.

1.4 Scope and Rationale

1.4.1 Scope

This thesis focuses on contract development software in contrast to packaged software. The field work was directed at finding the requirements in early design and connecting them to stakeholder goals. Changes in the environment that were either present (e.g., business model change) or artificially induced (e.g., viewpoint shifts) were supposed to impact the agreement with and priority of the requirements. The research goals were to establish a goal-oriented theory of requirements change, explaining why stakeholders want to change which requirements. Moreover, I wanted to verify these claims with empirical methods that came closest to a laboratory setting (i.e. field experiments). The research questions not only pertained to why requirements change but also how to create (new) measurement tools for investigating requirements change.

Within these boundaries, *requirements change* was regarded as the general tendency that the stakeholders' wishes for and demands on a system evolve. This can come about in many forms. Stakeholders may want to move from an existing into a new situation, for example, from a paper to a paperless office. Therefore, they implicitly or explicitly put up a wish list with *requirements*. In itself, this requirements list is a call for change. The features of the current system do not (completely) fulfill the stakeholders' needs any more. Another form of requirements change is that stakeholders add, delete, or adapt requirements during the off-line design of a system. The requirements list is not frozen yet, parts may still be under negotiation, and stakeholders can put up *change requirements*. Once the requirements list is fixed, the more technical development commences. Nevertheless, stakeholders may feel that certain aspects are not right yet. A bulky bug report, for example, may lead to filing a *change request*, which often comes in the shape of an electronic form.

To anticipate change requirements and change requests, the aim is to better understand the general phenomenon of requirements change. The challenge is to find the sources that predicate requirements change. My claim is that changes in requirements are directed by changes in the goals stakeholders have and to a lesser extent, the work processes they put to use to reach those goals, unless optimizing those processes becomes a goal in itself. The present book investigates what type of goals is responsible for changes in what type of requirements.

The broader context in which I worked was user-centered design (Chapter 2). With regard to RE, I worked from a goal-driven (e.g., Anton, Cracken, & Potts, 1994) and viewpoints-oriented approach (e.g., Kotonya & Sommerville,

1992). Another angle to the problem, I derived from emotion psychology (Frijda, 1986). The rationale of my work I outline next.

1.4.2 Rationale

Although at the surface, concrete manifestations of stakeholder needs may change over time, I believe that the more important needs or source concerns (cf. Frijda, 1986) of stakeholders probably do not. It is acclaimed that "goals are generally more stable than the requirements to achieve them" (Van Lamsweerde & Letier, 2000). Moreover, the higher-level a goal is (e.g., a strategic business goal), the more stable the respective requirements will be (Anton et al., 1994; Alves & Finkelstein, 2003). Therefore, I believe that the reasons for requirements change should first of all be sought in a change (in the weight) of lower-level goals, such as improving a work process (i.e. higher efficiency) or making cheaper products (i.e. making less costs).

My first angle was to study requirements change from a business perspective. Changes in the business model direct the requirements on a system (e.g., Lowe, 2003). However, writing or rewriting a business plan starts off with setting the personal goals that the business should achieve (e.g., Thayer, 1996; Palmer, 1999). "An entrepreneur's personal and business goals are inextricably linked" (Bhide, 1996), "… entrepreneurs build their businesses to fulfill personal goals" (ibid.), and "Attaining your business goals is the means of fulfilling your personal goals, not the end" (Thayer, 1996). In requirements engineering also, business and personal goals are analyzed as related entities (e.g., Loucopoulos & Kavakli, 1995). In other words, requirements change can be directed by changes in lower-level business as well as lower-level personal goals.

Much of what this book offers is based on empirical work. To gather the requirements, techniques were applied from ethnography (Jordan, 1996) and Groupware Task Analysis (Van der Veer et al., 1996). In most cases, participatory observations, document analysis, and interviews were enabled by arranging an internship with the organization or by involving a co-worker from the respective ICT department. The type of systems under investigation were a logistic warehouse management system (LWMS) at a provincial governance institution, a capacity management system (CMS) at the Dutch police force, an e-learning system (Didactor) at our university and at a financial buy-out of IBM, commercial off-the-shelf (COTS) computers, a tactile Braille mouse and related software, 25 financial systems at a Dutch multinational bank, and operation room equipment. Various informal field studies were necessary to gain insight into the organizational structure, the interactive systems under development, the business model, and the stakeholders' goals, concerns, and work processes. Requirements, goals, etc. gathered during the initial research phases were used as the contents of the questionnaires developed for requirements validation. The validation provided the data to confront my hypotheses with

and served as feedback for the IT practitioners to 'check the requirements with the stakeholders' (cf. Section 1.2).

All these studies, except one, concentrated on the conversion from an undesired existing situation to a desired future situation. The aim was to understand the occurrence of change requirements while the system was developed. The study on the multinational bank investigated the satisfaction with the current systems to understand better the incidence of change requests.

One means of investigating requirements change was the use of rank order tests on requirements lists (Chapter 3). I wanted to see whether and to what extent differences in requirements prioritization could be established under conditions of personal or business goal change. My other instruments of investigation (Chapter 4) were so called structured questionnaires (Dillman, 1999; Oosterveld, 1996; applied in Van der Raadt, Hoorn, & Van Vliet, 2005). The advantage over single-item surveys is that they control for interpretation ambiguities and measurement error (cf. Section 1.2). The core of the different variants of the Requirements Engineering questionnaire, RE*quest*, that I developed, featured a so called faceted scale (Guttman, 1965). The items on this scale systematically and explicitly connected a goal to a requirement. Each item, moreover, stated the positive or negative expectancy of using the requirement in the future situation. Stakeholders could rate these items for agreement (agree-disagree). The rationale was that variability in the level of agreement to requirements or goals indicated a certain degree of conflict among the stakeholders about the status of the requirement (should it be in or out?) or the goal. Consequently, the more variability in the level of agreement to a requirement could be measured, the more conflict this requirement probably raised among the stakeholders and conflict is a source for requirements change (Alves & Finkelstein, 2003). Put more technically, variance in the level of agreement to stakeholders' lower-level business and personal goals should explain the greater part of variance in agreement to requirements.

1.5 Overview of the Book

The chapters in this book are based on a number of papers presented at various conferences and workshops and published in conference proceedings and scientific journals. Each chapter will refer to this work, most of which is available through the ACM Digital Library (www.acm.org), my personal Web site, or the CD-ROM included in the book (Hoorn, 2005b).

Chapter 2 provides the theoretical background of the empirical work. This chapter discusses the Model of Requirements Change (MoRC) (Hoorn & Van der Veer, 2003a; 2003b) and how it evolved into the CoStaR model (Change of Stakeholder Requirements). The second part of Chapter 2 is transitional in that it drafts the methodological difficulties in conducting field experiments (cf. Shaw, 2002).

Chapter 3 reports on an empirical study at the Dutch police force (Hoorn, Breuker, & Kok, 2006). Here, police officers who were novice users to a Ca-

pacity Management System rated the importance of requirements from the point of view of the business and from their personal point of view. The data suggested that when presented from a business perspective, stakeholders hardly could imagine how the future system affected their personal work situation. I named this situation the 'requirements-analysis rift.' To avoid complaints after software implementation, analysts should relate requirements to (i.e. personal) goals.

This study was supplied with two rank order studies on requirements prioritization of an e-learning system (Hoorn & Breuker, 2005; Hoorn, Breuker, & Kok, 2006). Stakeholders put priority scores – supposedly indicating importance – to requirements from a business and a personal point of view. The dominance of personal over business goals again came to the fore. Changes in requirements priorities were significantly larger when personal goals changed than when business goals changed although the list of requirements was the same in both situations.

Chapter 4 discusses four related studies into the sources of requirements change (Hoorn et al., 2005; 2006). One study was conducted at a provincial governance institution, one with the said police officers, one with expert users assembling a commercial off-the-shelf computer from a predefined list of features, and the final study let blind pupils handle a Braille mouse. Results showed that the requirements most vulnerable to change were positive requirements that were connected to negative goal states and negative requirements connected to positive goal states. I coined this finding the 'goals-to-requirements chiasm.'

Chapter 5 introduces the model called Stakeholder Logistics (Hoorn, 2005a), which zooms in on user satisfaction after the system has been implemented and used. This model is validated using data from 25 financial systems at a multinational bank in The Netherlands. In addition, the research during the development of a progress monitor in the operation rooms of an academic medical center further explained the relation between efficiency-requirements, usability, and satisfaction. The results indicated that the major concern of both the bank and medical people was the efficiency of their machines rather than the usability. This concern was not so much about the (lack of) speed of their machines but how to keep them error free. In addition, these stakeholders saw human efficiency as completely separated from machine efficiency, which from a computer-human interaction viewpoint is curious to say the least.

Chapter 6 fuses all results into a new version of the CoStaR model with extensions to Stakeholder Logistics as they emerge from the empirical studies.

References

Alves, C., & Finkelstein, A. (2003). Investigating conflicts in COTS decision-making. *International Journal of Software Engineering and Knowledge Engineering, 13*, 1-21.

Anton, A. I., Cracken, W., Potts, C. (1994). Goal decomposition and scenario analysis in business process reengineering. *Proceedings of CAISE'94*, Utrecht, Netherlands.

Beyer, H. R., & Holzblatt, K. (1995). Apprenticing with the customer. *Communications of the ACM, 38, 5,* 45-52.

Bhide, A. V. (1996). The questions every entrepreneur must answer. *Harvard Business Review, 74, 6,* 120-130.

Boehm, B. W. (1981). *Software Engineering Economics.* Prentice-Hall.

Boehm, B., & Basili, V. R. (2001). Software defect reduction top 10 list. *IEEE Computer, 34, 1,* 135-137.

Boehm, B. W., & Pappacio, P. N. (1988) Understanding and controlling software costs. *IEEE Transactions on Software Engineering, 14,* 1462-1477.

Brooks, F. P. (1987). No silver bullet: Essence and accidents of software engineering. *IEEE Computer, 20, 4,* 10-19.

Checkland, P. (1999). *Systems Thinking, Systems Practice.* New York, NY: Wiley.

Coakes, E., & Elliman, T. Focus issue on legacy information systems and business process engineering: The role of stakeholders in managing change. *Communications of the AIS, Volume 2, 1, 4* (July 1999), 1-30.

Dillman, D. A. (1999). *Mail and Internet Surveys: The Tailored Design Method.* New York, NY: Wiley.

Edberg, D., & Olfman, L. (2001). Organizational learning through the process of enhancing information systems. *34th Hawaii International Conference on System Sciences,* IEEE, http://csdl.computer.org/comp/proceedings/hicss/2001/0981/04/09814025.pdf.

Frijda, N. H. (1986). *The Emotions.* New York: Cambridge University.

Guttman, L. (1965). A faceted definition of intelligence. *Studies in Psychology, 14,* 166-181.

Hoorn, J. F. (2005a). Sources of requirements change. A goal and viewpoints-driven explanation. *SIGCHI.NL* The Hague, Netherlands, October 13, 2005. New York: ACM Press.

Hoorn, J. F. (2005b). *Sources of Requirements Change or Why Requirements Engineering is such a Tough Job. Report SenterNovem IOP MMI99009* [CD]. Amsterdam: Breckfield Hall.

Hoorn, J. F., & Breuker, M. E. (2005). Panta rho: Personal before business in changing requirements priorities. In J. Ralyté, P. J. Ågerfalk, & N. Kraiem (Eds.), *Proceedings of the First International Workshop on Situational Requirements Engineering Processes: Methods, Techniques, and Tools to Support Situation-Specific Requirements Engineering Processes (SREP'05), Paris, France, August 29-30, 2005, In Conjunction with the Thirteenth IEEE Requirements Engineering Conference (RE'05)* (pp. 152-162). Ireland: University of Limerick.

Hoorn, J. F., Breuker, M. E., & Kok, E. (2006). Shifts in foci and priorities. Different relevance of requirements to changing goals yields conflicting prioritizations and is viewpoint-dependent. *Software Process: Improvement and Practice.* In press.

Hoorn, J. F., Konijn, E. A., Van Vliet, H., & Van der Veer, G. (2005). Goal-oriented RE for handling change requirements: An explanation of what stakeholders try to avoid and what they try to achieve. In K. Cox, E. Dubois, Y. Pigneur, J. Verner, S. Bleistein, A. M. Davis, & R. Wieringa (Eds.), *Proceedings of the International Workshop on Requirements Engineering for Business Need and IT Alignment* (REBNITA'05), Paris, France, August 29-30, 2005, In Conjunction with the Thir-

teenth IEEE Requirements Engineering Conference (RE'05) (pp. 123-132). Australia: University of New South Wales Press.

Hoorn, J. F., Konijn, E. A., Van Vliet, H., & Van der Veer, G. (2006). Requirements change: Fears dictate the must haves; desires the won't haves. *Journal of Systems and Software.* In press.

Hoorn, J. F., & Van der Veer, G. C. (2003a). Requirements analysis and task design in dynamic environments. In D. Harris, V. Duffy, M. Smith, & C. Stephanidis (Eds.), *Human-centred computing: Cognitive, social, and ergonomic aspects. Volume 3 of the proceedings of HCI International 2003*, Hersonissos, Crete, Greece (pp. 472-476). Mahwah, NJ: Erlbaum.

Hoorn, J. F., & Van der Veer, G. C. (2003b). Requirements engineering for groupware to manage business process dynamics and stakeholders' mindset. In G.C. van der Veer, & J.F. Hoorn (Eds.), *Proceedings of the 9th European Conference on Cognitive Science Approaches to Process Control: Cognition and Collaboration - Distributed Cognition in Complex Processes 2003*, Amsterdam, Netherlands (pp. 153-160).

Hsia, P., Davis, A., & Kung, D. (1993). Status report: Requirements engineering. *IEEE Software 10, 6*, 75-79.

Jordan, B. (1996). Ethnographic workplace study and CSCW. In D. Shapiro, M. J. Tauber, & R. Traunmüller (Eds.). *The Design of Computer Supported Collaborative Work and Groupware Systems.* Elsevier Science.

Kotonya, G., & Sommerville, I. (1992). Viewpoints for requirements definition. *Software Engineering Journal, 7,* 6, 375-387.

Kovitz, B. (2003). Hidden skills that support phased and agile requirements engineering. *Requirements Engineering, 8,* 135-141.

Leffingwell, D., & Widrig, D. (2003). *Managing Software Requirements: A Use Case Approach.* Boston, MA: Addison Wesley Professional.

Loucopoulos, P., & Kavakli, E. (1995). Enterprise modeling and the teleological approach to requirements engineering. *International Journal of Intelligent and Cooperative Systems, 4, 1,* 45-79.

Lowe, D. (2003). Web system requirements: An overview. *Requirements Engineering, 8, 2,* 102-113.

McConnell, S. (1993). *Code Complete: A Practical Handbook of Software Construction.* Redmond, WA: Microsoft Press.

Oosterveld, P. (1996). *Questionnaire Design Methods.* Nijmegen: Berkhout.

Palmer, R. (1999). *Dairy Modernization Planning Guide - Consolidate Personal Goals To Develop Business Goals.* DairyBusiness Communications A Multi Ag Media Company - University of Wisconsin-Madison. Retrieved Sept. 8, 2005 from http://www.dairybusiness.com/midwest/May2000/goals.htm

Shaw, M. (2002). What makes good research in Software Engineering? *International Journal of Software Tools for Technology Transfer, 4, 1,* 1-7.

Simonsen, J., & Kensing, F. (1997). Using ethnography in contextual design. *Communications of the ACM, 40, 7,* 82-88.

Thayer, C. (1996). *Setting Up Your Own Business: Balancing Business and Personal Life.* University of Nebraska-Lincoln. Retrieved Sept. 8, 2005 from http://ianrpubs.unl.edu/consumered/nf271.htm

Van der Raadt, B., Hoorn, J. F., Van Vliet, J. C. (2005). Alignment and maturity are siblings in architecture assessment. In O. Pastor, & J. Falcão e Cunha (Eds.), *CAiSE 2005*, LNCS 3520 (pp. 357-371). Berlin, Heidelberg, Germany: Springer

Van der Veer, G. C., Lenting, B., & Bergevoet, B. (1996). GTA: Groupware Task Analysis – modeling complexity. *Acta Psychologica, 91, 3,* 297-322.

Van Lamsweerde, A., and Letier, E. (2000). Handling obstacles in goal-oriented requirements engineering. *Software Engineering, 26, 10,* 978-1005.

Van Vliet, H. (2000). *Software Engineering. Principles and Practice.* New York, NY: Wiley.

Viller, S., & Sommerville, I. (2000). Ethnographically informed analysis for software engineers. *International Journal of Human-Computer Studies, 53,* 169-196.

Wiegers, K. E. (2003). *Software Requirements.* Redmond, WA: Microsoft Press.

Wikiquote (2005). *Charles Darwin.* Retrieved August 28, 2005 from http://en.wikiquote.org/ wiki/Charles_Darwin

WorkSmart™Guide (2005). *Brightcode's process to deliver On-Time, On-Budget, On-Target Information Technology.* Retrieved July 28, 2005 from http://www.brightcode.com/ downloads/WorkSmartGuide.pdf

2 Theory

Abstract

In this chapter,[1] the broader theoretical background is sketched of the empirical work that this book presents. On a meta-level, the user-centered design framework in which this work is performed is DUTCH (Designing for Users and Tasks from Concepts to Handles). Here, task analysis and (detail) design of the User Virtual Machine are informed by several sub disciplines to understand the stakeholders' culture, organization forms, psychology, and financial situation (Section 2.1). The DUTCH approach and in particular its task analysis aspects (Section 2.2) are enriched with concepts from business modeling (Section 2.3). This extended DUTCH model (Section 2.4) is then compared to more general approaches in RE (Section 2.5). Section 2.6 argues that business and task modeling is probably not sufficient to account for requirements change. Motivational aspects and personal goals (Section 2.7) can explain a change as well. Regarding business processes, Section 2.8 zooms in on speed and accuracy aspects of processes because this mechanism generally applies to business, computer, and human processing. Section 2.8 assumes that changes in these process dimensions may strongly affect what is demanded of a new system. Finally, a number of concepts is integrated into a model that claims to explain requirements change (Section 2.9). This model is then refined into the CoStaR model (Change of Stakeholder Requirements) (Section 2.10). Section 2.11 is transitional in that it discusses the pitfalls of doing research in real business settings so that the reader is better prepared to evaluate the empirical chapters that go subsequent.

Keywords: DUTCH, GTA, Business models, ISAC, Relevance, Valence, Personal goals, Speed-accuracy trade-off

2.1 General Background: DUTCH Design

System designers and requirements engineers often choose a framework to elicit, analyze, and refine the requirements list. Such a framework helps to focus on important aspects in the organizational and task environment. The DUTCH design approach (Designing for Users and Tasks from Concepts to Handles) is a general method to design complex interactive systems (Van Welie & Van der Veer, 2000). DUTCH takes into account, for instance, the work situation, the task world, and the system stakeholders in a dynamic way,

[1] This chapter is based on Hoorn and Van der Veer (2003a; 2003b), Hoorn (2004, Tech. Rep. [CD]), and Hoorn et al. (2004).

so that the approach renders possibilities to get to some form of requirements management. The analytical part of the DUTCH approach (Figure 1) is called Groupware Task Analysis (GTA) (Van Welie, 2001; Van der Veer, Lenting, & Bergevoet, 1996). Instead of merely analyzing the individual at work, GTA also recognizes the impact of the work group, business, or organization (e.g., Chisăliță et al., 2005). To allow for more individual, work, and situational aspects, two task models are discerned. One describes the state of affairs of the task environment on the work floor (Task Model 1) and the other envisages the desired situation after the system has been built and implemented (Task Model 2). Because Task Model 2 (TM2) describes a more-or-less ideal situation, it is limited by practical constraints such as technical possibilities, the client's budget, and legislative directives. During and after analysis, specifications of the requirements are prepared, guiding detailed decisions about 'what the stakeholder actually wants.' Once the system is implemented, however, stakeholders do not clearly differentiate between hardware and software. Before implementation, then, this so-called 'User Virtual Machine' (UVM) needs to be designed. In the mind of the stakeholder, functions, dialogs, and representations are pragmatically fused, irrespective of their physical or conceptual origin (Tauber, 1988; Van der Veer & Van Vliet, 2001). Designers should anticipate this in their detail design. A special technique for specification notation is NUAN (New User Action Notation, see Van Welie, 2001, chap. 5.7). Together with TM2, NUAN provides about the same information as a requirements specification does (Van der Veer & Van Vliet, 2001). Evaluation and validation of the system can be done by playing scenarios, doing simulations, and developing prototypes and mock-ups that are used in real life environments. It is here that requirements change can be detected, which may be quite late in the project's time line.

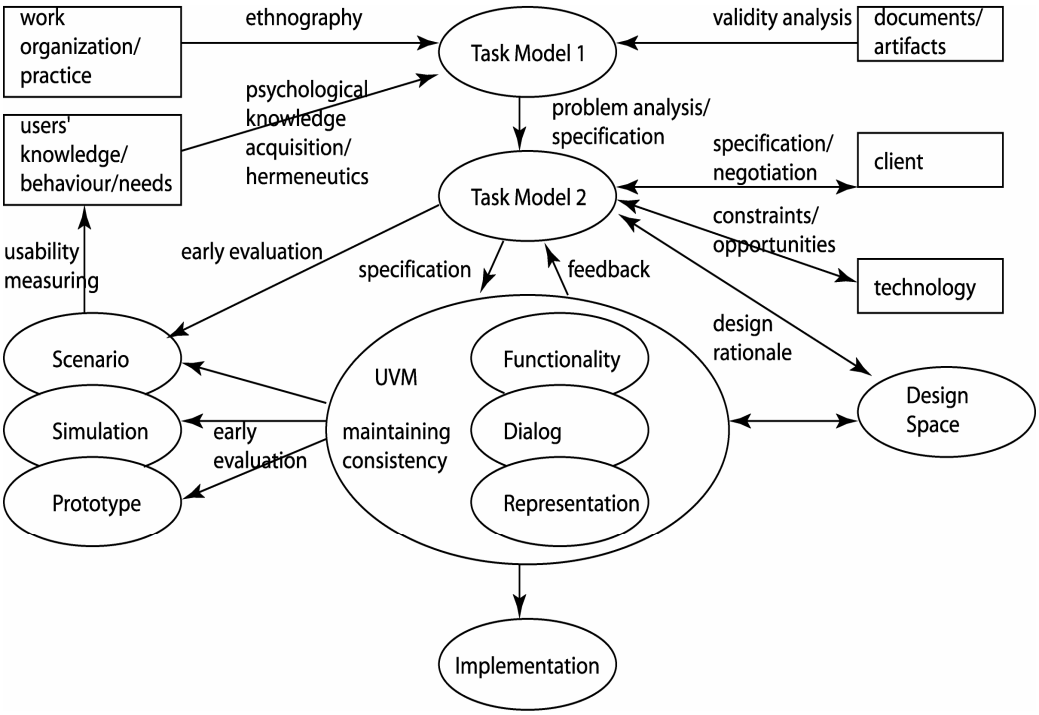


Figure 1: The DUTCH design approach (Van Welie, 2001, p. 3).

The present research is concerned with the upper part of Figure 1; analysis of the organization, work practice, and stakeholders' needs through ethnography and psychology. Also task and document analysis is performed to acquire an initial set of requirements. Yet, the work stops when the UVM should be created. In other words, my work is not about detail design, making prototypes and mock-ups, or implementation.

2.2 Groupware Task Analysis

Within the DUTCH approach, a central activity is task analysis. To study the elements that make up a task world, Groupware Task Analysis (GTA) proposes a usable ontology (e.g., Van Welie, 2001, pp. 43-44). A task is an activity performed by people or software (both referred to as 'agents') to achieve a particular goal (Figure 2). A task typically makes a change in the task world and takes time to complete. Complex tasks can be decomposed into smaller sub tasks. Tasks are performed in a specific order and finishing one task can trigger another.

A task can be invoked by an event that happens in the task world, for example, handling a new delivery. A goal is a desired state in the task world or a desired state of the system. A goal can be achieved by one or more tasks and may have sub goals. In the Van Welie task model, a goal can be both a personal goal or a business goal.

A meaningful compilation of tasks executed by (a group of) agents is called a role. The role becomes meaningful in view of a clear goal or when it

differentiates groups of agents or stakeholders (e.g., managers vs. workers). Roles can be hierarchically arranged and are accountable for the tasks they imply.

An object refers to a physical or non-physical entity. The latter could be thought of as, for example, messages, signatures, gestures, and stories. Objects have attributes consisting of attribute-name and value pairs. Actions stipulate what can be done with an object (e.g., an e-mail message): To move, transmit, turn on, etc. Objects can belong to a type hierarchy and can be contained by other objects.

An agent is an entity that is believed to be active in the task world. It may be a human operator as well as a software component that performs a task. Agents can be one individual but also a category of individuals with their characteristics.

An event is a change in the task world's condition at a given point in time. The change may reflect shifts of attribute values of internal concepts such as object, task, agent or role. They could also reflect changes of external concepts such as political climate or business situation. Events affect the order of task performance by triggering tasks. In using the object-oriented Unified Modeling Language (UML), the work by Döring, Dörfel, and Distelmaier (2001) on military air traffic illustrates that the GTA ontology is sufficient to represent task domains.

In my research, GTA was used to study the stakeholders at the work floor. For purposes of theoretical modeling, the aspect of 'events' was taken to a more general level, that events – whether personal or business – could incite requirements change. As said, in GTA, a goal can be personal or business. However, to study the different effects of both, I decided to split the goals up into those goals typical for the business and those typical for personal concerns. Thus, within the task world, the goal to achieve with a system can be personal or business but this input comes from a business model or a 'personal model.' As 'roles' I chose to study the differences between managers vs. workers (i.e. Chapter 3 and 4). How the goals differentiated these groups was laid down in so-called 'viewpoints' (see Chapter 3).

Within the DUTCH design approach, GTA provides for changes in a work situation by means of task models. Events, tasks, and agents affect their environment and NUAN models in relation with TM2 resemble the requirements specification. However, information about business models is needed to improve this resemblance.

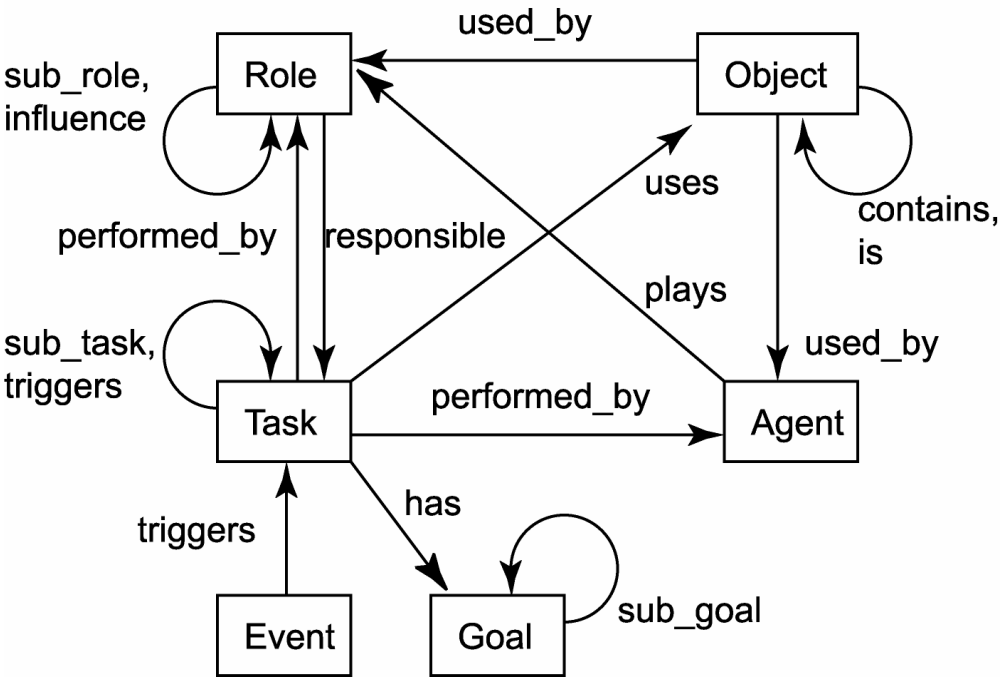


Figure 2: Van Welie task model.

2.3 Business Situations

A well-known requirements engineering method is Information Systems Work and Analysis of Changes (ISAC) (e.g., Lundberg, Goldkuhl, & Nissen, 1981; Wieringa, 1996). Although a bit dated, ISAC uses activity models that represent a current or desired business situation in quite a clear way (Figure 3).

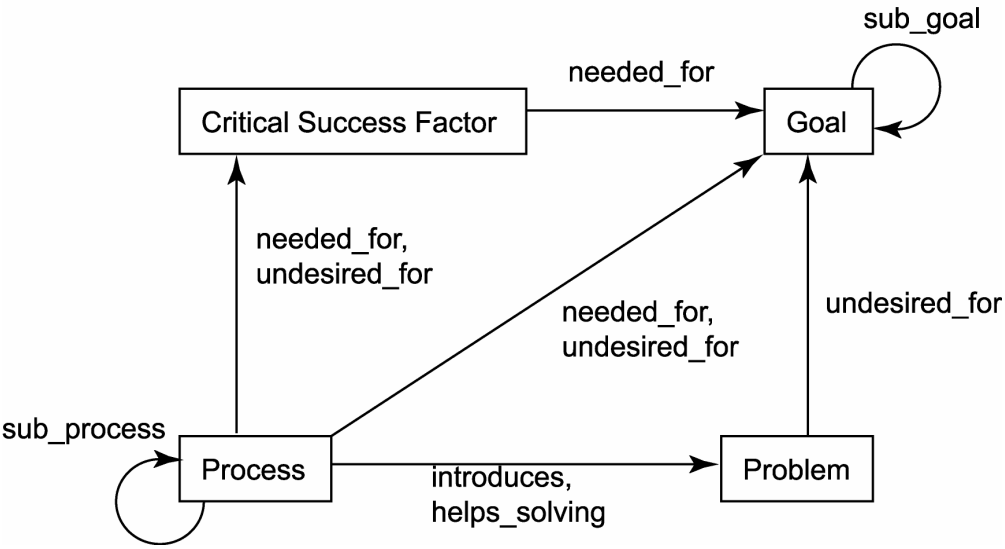


Figure 3: ISAC activity model for current or desired business situations.

In Figure 3, a goal is a desired situation of the business somewhere in the future. A business goal is not the desired result of a fixed set of processes, but a desired state of (part of) the business in the future. It may be difficult to relate this type of goal to particular processes. Here also, goals can be divided into sub goals if desired. A business problem may cause dissatisfaction with the current situation. Problems may be phenomena of which the management believes that they hinder achieving a goal. A Critical Success Factor (CSF) is an area in which, according to the management, the business must perform well to achieve certain business goals. A CSF may be the available budget or human resources. A business process is an activity that transforms an input into an output. Processes may be split into sub processes. The relations between the processes and goals can be positive (process needed_for) or negative (process undesired_for).

Several aspects of ISAC could be used to determine the input of the task model (previous section) and to advance a model of requirements change. The goal of a task could be a business goal, for instance, an increase in the number of sales events. Reaching this goal could be done by changing the way a business transaction is performed. Placing an order through the telephone could be replaced by automated Web transaction. This change in the business process typically affects the task execution and finally, how well the goal is accomplished. Thus, a model of requirements change should account for business goals as well as processes because both aspects can give rise to a request for a change. In addition, expectations of the stakeholders also may impact the urge to change. Envisioned problems are perceived as undesirable for goal accomplishment and the CSFs are seen as needed or desired to reach a goal. Likewise, certain processes may be desirable (e.g., quick-fix solutions) or unwished for (e.g., doing many work-arounds) to achieve a goal. In Section 2.6, I argue how the pair needed_for – undesired_for link up with personal models and human affect.

2.4 Integrating Business Models with Task Analysis

To arrive at requirements management that on the one hand can account for requirements change and on the other provides a specification of the requirements that are agreed-upon by the stakeholders, the ISAC activity model of Figure 3 should be integrated with the GTA in DUTCH. The sediment of such an attempt is the adapted version of the DUTCH approach in Figure 4. For simplicity, all evaluation activities have been omitted, as well as the specification of the UVM.

The task hierarchy of Business Model 1 (BM1) corresponds to the task or process hierarchy in TM1. The task hierarchy of Business Model 2 (BM2) corresponds to the task or process hierarchy in TM2. BM1 relates business requirements to TM1 (goals, CSFs, problems). The single-headed arrow from TM1 (work floor) to BM1 (executive office) could actually be double-headed

in that the work floor should have an understanding of the business plans as well.

BM1 and TM1 are analyzed, leading to TM2. The task hierarchy of TM2 is used to build BM2, which relates the business requirements to the desired situation. BM2 illustrates how business goals, CSFs, and problems affect TM2 and serves as a justification of TM2 to that extent.

New technology can strongly affect the new business model (e.g., Zlatev et al., 2004). Therefore, a double-headed arrow between BM2 and TM2 should be common practice but it demands a client who cooperates strongly with the management to help design the future business model. In this respect, Zlatev et al. (2004) offer an elegant way to apply goal-oriented RE to create new business models from value patterns.

Once there is a specification of the User Virtual Machine with NUAN models, the implementation can start. As already mentioned, the NUAN models and TM2 together provide about the same information as a requirement specification does.

This similarity may be enhanced by taking business requirements into account. Business goals may have a strong impact on task hierarchy. Business goals probably affect the task hierarchy on a higher level (closer to the root) than user requirements do. If such a business goal is at issue, it limits the use of a detailed investigation of the current tasks at a lower level in TM1. For this reason, a design team should not wait with obtaining business requirements until TM1 is finished.

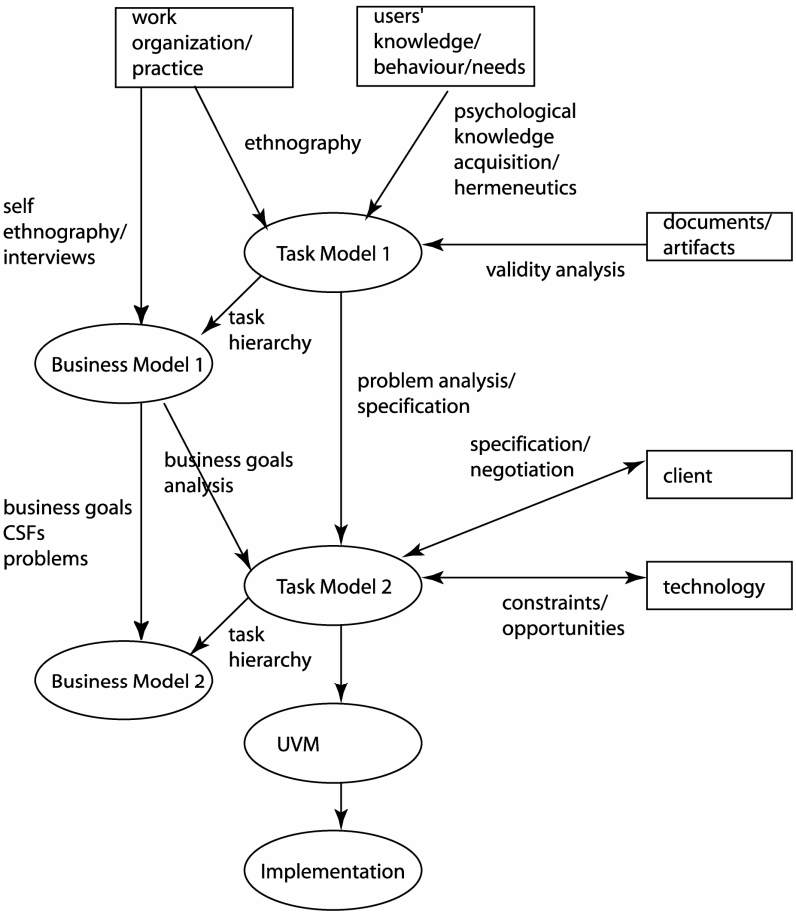


Figure 4: The DUTCH process model with business model extensions.

2.5 The 'Engineering' of Requirements

In between Business-plus-Task Model 1 (BTM1) and Business-plus-Task Model 2 (BTM2) (Figure 4), the actual 'engineering' or construing of requirements takes place. On the input side of the general requirements method depicted in Figure 5 (Kotonya & Sommerville, 1998) (left column), part of the system requirements is revealed through information from stakeholders, business and task analysis (e.g., Sebillotte, 1995), system documents, domain knowledge, market studies, ethnography, questionnaire studies, etc. (cf. BTM1). In the process of requirements engineering (Figure 5, middle box), this is the stage of requirements elicitation.

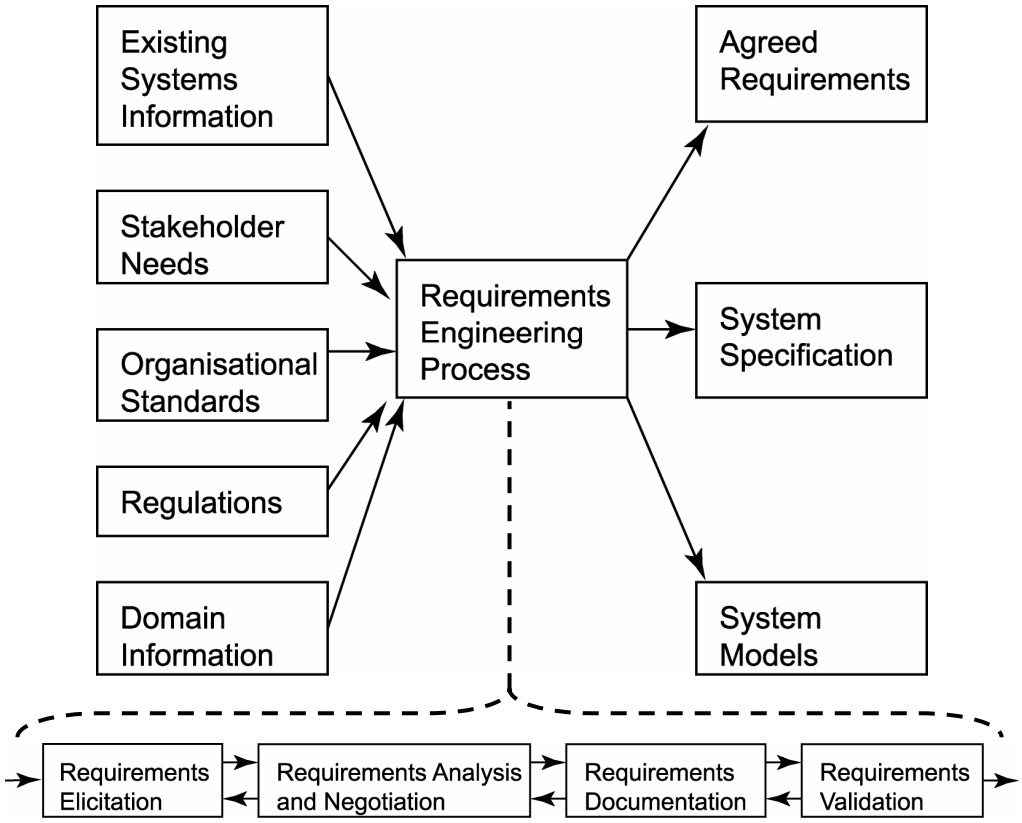


Figure 5: Inputs and outputs of the requirements engineering process.

The stages identified in the requirements engineering process (Figure 5, dashed) are not structured as nicely as appears from the figure but the activities do take place more-or-less in the given order.

The second stage in the dashed part of Figure 5 indicates that requirements are analyzed in detail. The different stakeholders negotiate requirements to avoid a conflicting system.

The third stage is documenting the agreed requirements at an appropriate level of detail and in a way that is understandable to all stakeholders. Finally there should be a thorough check for consistency and completeness of the requirements (validation).

Because of the sometimes-conflicting requirements, next to the stages in the requirements engineering process there should be some kind of requirements management in which changes to the requirements of a system are controlled (e.g., Natt och Dag et al., 2005).

Regarding the output (Figure 5, right column), this thesis does not focus so much on system specifications or system models but rather on the agreed-upon requirements (cf. BTM2 and UVM); the area where the mismatches between what was requested and what can be or is delivered come to light.

However, interlacing the general requirements engineering model by Kotonya and Sommerville (1998) with DUTCH and DUTCH with business

models is probably not enough to account for (dis)agreement to requirements, and hence, for requirements change. As stated in Section 2.2, the goal of a computer task also could be a personal goal.

2.6 Task-based Worlds – Task-based Emotions

Next, I want to turn to the motivational aspects that accompany the achievement of goals, the way people personally deal with problems or how they respond to market events. Probably, agreement or disagreement to a requirement is not only fed by the degree to which that requirement serves a goal but also by the stakeholders' affective experiences with that goal or requirement. A judgment whether a requirement should be on or off the system is probably susceptible to emotional biases.

The two fundamental factors that regulate affective responses are relevance and valence. The work of, among others, Arnold (1960), De Sousa (1987), Lazarus (1982), and Sartre (1936), inspired Frijda (1986, p. 494, p. 463) to propose a functional theory of emotion. Frijda states that after an event (in the surroundings or in imagination) has taken on meaning for a person, the features in the situation are judged for their usefulness or *relevance* to potentially satisfy or harm one's concerns, goals, or motives. If the situation or event appears to be irrelevant, the emotion process stops: Without relevance, no emotions occur. Emotions result from the match or mismatch between events and concerns. Relevance stimulates the intensity of emotions, whether they are positive or negative. If an event touches upon an important goal (e.g., not getting fired), the accompanying emotions are strong.

The direction of emotions (positive, neutral, or negative) largely evolves from what is called *valence*. Positive emotions result from events that promise a match, the actual or expected satisfaction of concerns. Negative emotions result from events that promise a mismatch, the actual or expected obstruction of realization of goals and concerns (Frijda, 1986, p. 277). Frijda (1986, p. 207) points out that valence refers to the implied outcome of the event or 'what you expect of it.' This is the intrinsic attractiveness or repulsiveness of a situation. Valence refers to the expected match or mismatch between an event and the potential gratification for or obstruction of one's concerns. People evaluate a situation for its possibilities or impossibilities to satisfy or harm a concern. Thus, valence reflects the establishment of positive and negative directions of affect, which can interact with relevance (strongly negative, weakly positive, etc).

Hence, emotions probably result also from the match or mismatch between requirements and goals. Positive emotions result from requirements that promise a match, the actual or expected satisfaction of goals. Negative emotions result from requirements that promise a mismatch, the actual or expected obstruction of realization of goals. The more important the goal that should be accomplished with the system, the more intense the emotions are. If automation puts a person out of a job, the direction of affect towards the requirements

will be strongly negative. It is not hard to see that such emotions bias the agreement to requirements. If someone suspects that the new work situation will harm his or her position, the disagreement to requirements is high. If an automated teller machine in a foreign country refuses your card, valence towards the machine will be negative and becomes stronger the less money you have in your pocket. Disagreement to the machine's features in the current task situation will be high and the call for a change into the direction of universal access in the future will be loud. Likewise, if a business process is 'needed_for' increasing the number of sales-events (Figure 3), the valence towards that process is positive compared to a process that is 'undesired_for' or obstructs reaching that objective.

Because most of the goals that should be achieved with a system are task related, the task-based emotion theory of Konijn (2000) may provide an interesting angle. In her empirical work, Konijn found that particular emotions such as joy, anger, challenge, boredom, and listlessness in professional stage actors were uniquely related to task performance. This type of emotions is similar to what everyday users experience while interacting with their computers, i.e. while playing games (cf. Computers as Theatre – Laurel, 1991). Task-emotion theory may show useful to explain certain aspects of user experience but this is beyond the scope of the present study.

2.7 Goal Types: Personal vs. Business, Egotistic vs. Altruistic

Althought the literature (e.g., Bertrand et al., 1998) emphasizes the importance of goal-driven RE, there might be many types of goals that explain requirements change. Business goals are among the most popular (e.g., Robertson & Robertson, 2004) but there are authors who believe (see next) that business goals are instantiations of more general human needs and that businesses typically are set up and changed to achieve *personal* goals.

Therefore, requirements change may be based not only on a focus shift between business goals but between personal goals as well. Cooper and Reimann (2003) indicate that interactive systems are more successful when they are designed not only from the point of view of the business but take the personal goals of the users into account as well. Price and Cybulski (2004) state that stakeholders "often value their personal goals over those of their employer" and that it is important to align requirements with personal goals to keep a software development project from failure. Some of the examples Cooper and Reimann (2003, pp. 55-74) provide of personal goals that a system should meet are feeling in control, being competent, being successful, feeling sure of yourself, or to get promoted. These goals go beyond the restricted task domain that GTA covers. Thus, the type of goals that should be studied in TM1 should go beyond the strict computer task or the business goals and look at personal motivational aspects of the workers or managers as well. If a data typist runs the risk of being fired because her job will be computerized, s/he will disagree strongly to the requirement of automated input.

Requirements change may also come from a shift within the business or personal goals. A classic business goal is maximizing profit (cf. Hirschheim & Klein, 1989: Functionalism paradigm). However, in the nonprofit sector the mission is more philanthropic (ibid: Neohumanism paradigm). Yet, both business types have commercial as well as ideological goals albeit in different numbers with different degrees of importance. A hard-boiled commercial business like Philips at least used to advertise 'Let's make things better.' Amazon.com wants to bring books to the world but should stay financially sound as well. The current call for market-driven organizations may lead the goals of a nonprofit organization from reaching break-even to making profit, focusing not only on the common good but also on being self supporting and maintaining business continuity. Because requirements are usually seen as refinements of goals (e.g., Darimont & Van Lamsweerde, 1996), shifts from public to organization centered goals may cause requirements change. Thus, if a nonprofit organization is privatized, altruistically-oriented goals shift to or are supplied with more egotistically-oriented goals. As a consequence, the service-oriented, open source application may be supplemented with a protected customer transaction system.

The same distinction between egotism and altruism can be made also in the personal goals of employees (whether laborers or managers). In life and work, people can be more selfish or more socially oriented. On the one hand, they work to improve their own circumstances (e.g., the pursuit of money, power, and replication possibilities, cf. Singer, 1996; Steensma, 1999). On the other, employees show loyalty to the group (e.g., Van Leeuwen, 2001) (a work team or a business) and may subscribe to the business goals. When external events occur, the trade-off between egotistic and altruistic goals may result in a different demand on the system requirements.

2.8 Aspects of Processing: Speed-accuracy Trade-off

There are many aspects of a business process that could be studied, for example, features vs. delivery or features vs. correctness. To advance a general theory of requirements change, however, I attempted to find a mechanism that could apply to both human, computer, and business processes and that was dynamic enough to be a candidate for the explanation of change. Moreover, I wanted to sharpen the conceptual understanding of the in business and computer science almost over-used term 'efficiency' and all this I could establish by focusing on the speed and accuracy aspects of executing processes.

Quite some computer science literature is based on optimizing processing speed or minimizing error repair.[2] Similar trends can be detected in business process optimization.[3] Faster transaction times and error reduction are benefi-

[2] The search <optimize computer "processing speed"> in Google Scholar Beta yielded 1,570 hits; <minimize "system error"> yielded 1,600 hits (August 18, 2005).

[3] The search <optimize speed "business process"> in Google Scholar Beta yielded 1,200 hits; <minimize error "business process"> yielded 1,390 hits (August 18, 2005).

cial to business processes as well. When asked to perform a task as well as possible, people will apply various strategies that may optimize speed, optimize accuracy, or a combination of the two. For this reason, comparing the performance of a number of users cannot be done on the basis of speed or accuracy alone, but both values need to be known (Usability Glossary, 2005). This renowned speed-accuracy trade-off may be responsible for changes in the business or work processes that a software project should support. When stakeholders feel that one system shell on top of the other will increase the error percentage or slow down processing speed, they may ask for a complete redesign of the operating system. To gain high speeds they may tolerate more errors or, in the case of a safety critical system such as a hospital device, stakeholders may give in on the speed aspect so to increase accuracy, for example, by adding control loops.

The typical software engineer or business analyst wants processes to be efficient, meaning that they are executed accurately while being fast. The counterpart, of course, is that inefficient processes are inaccurate (error prone) while being slow. In between, there are sub optimal combinations, such as processes that are accurate but slow (e.g., safety critical systems) or inaccurate but fast (cf. the example of an image transmission control mechanism next):

> … a notion of quality of service can be used which allows the user to control the error margin: "guaranteed" will try to decrease the error to nearly zero at the cost of a long pre-processing time, "best effort" will use the multi-pass method with an error threshold in order to deliver a moderate prediction error, and "no guarantee" will deliver inaccurate but fast results using the two step method. (Rauschenbach & Schumann, 1997)

The above example also illustrates that what the right (combination of) speed and accuracy should be for a system is a matter of perception, that is, in view of the prevalence of the goal that is affected by the process. Managers of a commercial business may insist on transaction speed perhaps at the cost of correct service. A machine operator may insist on slower processing times because the machine makes him or her work too hard or let him/her make too many errors. The model of requirements change presented next will treat the speed-accuracy trade-off in executing processes as one source of requirements change.

2.9 Outline: Model of Requirements Change (MoRC)

To arrive at the first layout of a model of requirements change, I used the following concepts. The idea that what a system will look like is based on the agreed-upon requirements stems from Kotonya and Sommerville (1998) (Figure 5). That the cause of all changes is rooted in events (whether business or personal) is derived from GTA (Figure 2). That business goals and processes contribute to goal achievement and thus, to the requirements definition, stems from ISAC (Figure 3). ISAC also provided the entrance to emotion psychol-

ogy (i.e. valence, Frijda, 1986) by emphasizing the importance of stakeholder expectations (needed_for, undesired_for). That a current and future situation should be taken into account is drawn from GTA within DUTCH (Figure 1, Figure 4).

In Figure 6, a first attempt is made to structure some of the assumptions and theoretical concepts that were explained in this chapter. Figure 6 forms the beginning of a path model in which causal relations and correlations among variables are represented. For the sake of simplicity, the model in Figure 6 is not complete and not designed how it should be formally (e.g., Rigdon, 1998). Therefore, it is provisional and only there to shape a first understanding of the framework in which the empirical work was performed.

Concepts in boxes with solid lines should be read as variables that can be observed in the outside world. An external event such as an aggressive take-over is an example of such a variable but also a filled-out change-request form is. Observable variables are also the requirements statements on a questionnaire that can be scored for agreement (Figure 6, solid boxes within the dashed boxes). Such statements can concern the profit a company wants to make, product quality, or the type of process that is required (e.g., quick and dirty).

Concepts or 'constructs' in ellipses are unobservable or latent variables. Together they form 'the stakeholder's mindset.' These variables typically are judgments that stakeholders make on the basis of external stimuli (e.g., events or system features). The variables within dashed boxes are trade-offs between goals and/or processes that contribute to the level of agreement to requirements. For example, if over time the trade-offs within the Goals box are changing, this directs different levels of agreement to requirements. Likewise, if processing speed is traded for accuracy or v.v., this will change the level of agreement to requirements.

In Figure 6, arrows depict the relations between the concepts or 'constructs.' The arrows represent the hypotheses of the model. Single-headed arrows describe dependencies (e.g., agreement to requirements depends on the relevance of goals). Double-headed arrows describe correlations (e.g., goals and processes are mutually connected).

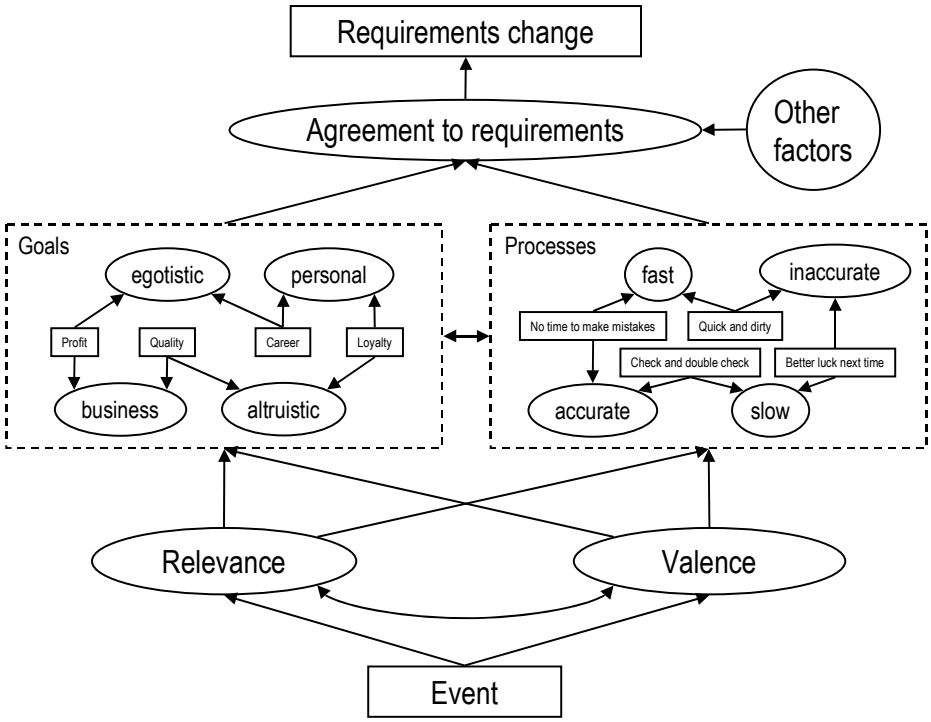


Figure 6: Model of Requirements Change (MoRC).

The representation of the MoRC in Figure 6 should be read from bottom to top. The MoRC (Hoorn & Van der Veer, 2003a; 2003b) assumes that requirements change is instigated by an event, for example, a technological breakthrough (cf. Zlatev et al., 2004). Stakeholders evaluate this event in terms of relevance (gravity, importance, urgency) and valence (positive, negative, or neutral expectations) with regard to their goals (whether business or personal) and processes. Goals can be business or personal and are egotistically or altruistically oriented. Maximizing profit is an example of a self-oriented, egotistic business goal. Producing high-quality products exemplifies a more altruistically-oriented business goal (i.e. pleasing the customer). Continuing or improving a career is an egotistic-personal goal. If someone is put out of a job due to intensive automation, this stakeholder will strongly object to the new technology. Being loyal to the work team or organization (cf. police or army) is an example of an altruistic-personal goal. If registering shifts and working hours in e-forms takes more time than the usual verbal reports, an employee may resist that new system feature because she has less time to help her colleagues, customers, or chief. This last example also illustrates the relation between goals and processes. The goal of helping colleagues is harmed via the time delay of filling out e-forms.

If the business situation is such that there is no time to make mistakes (e.g., in an operation room), technology should guarantee fast and accurate processing. When a system is in this state, it probably is judged ‘efficient.’ If

the goal is to make money without caring too much about quality, the supporting business process is allowed to be quick and dirty, that is, possibly inaccurate but necessarily fast. If the technology should support air traffic control or nuclear power systems, the demand of safety and double check-ups goes before a loss of speed. That is, speed can be traded for accuracy. When someone surfs the Web without a specific goal in mind (Bhulai & Van der Veer, 2005), this 'wandering around' is time-consuming (slow) and is allowed to be inaccurate (better luck next time). Such a search process may be deemed 'inefficient' but without the negative connotation. Commercial enterprises may want to attract such traffic to their Web sites, because these potential customers may just stumble upon something of their desire.

The dashed boxes in Figure 6 contain the trade-offs within goals and within processes. A new market event (e.g., deregulation or privatization) may demand to trade safety for money or accuracy for speed. This changes the level of agreement to the requirements on the supporting systems, eventually leading to a change request. These trade-offs are regulated by shifts in relevance and valence. In view of emerging developments (e.g., dense competition), features on an existing system or requirements on a new system may raise less agreement because certain goals (e.g., safety) become less relevant whereas other goals become more relevant (e.g., make money). Similarly, the direction of valence may change. A non-profit organization may have a negative attitude to charging money for services but once privatized, charging money may be regarded positive, thus demanding a customer transaction system. The intricate dynamics of requirements change, according to the MoRC, is that the shifts between business and personal, between egotistic and altruistic, and between speed and accuracy may occur simultaneously. This is because the relevance and valence of these variables change when a new event occurs.

Probably, not only goals and processes lead to variability in agreement to requirements. For instance, past experiences, negotiation skills, and power relationships may contribute to how much a stakeholder agrees to the proposed system features. However, these factors fall outside the scope of this study and are accounted for as unexplained variance caused by 'other factors.' The MoRC expects that when stakeholders generally agree to requirements, requirements change remains absent (ceiling effect of agreement). When in general, stakeholders disagree to a requirement, requirements change will be absent as well (ceiling effect of disagreement). Variability in the level of agreement only occurs when stakeholders show mixes of agreement and disagreement with requirements.

2.10 Adaptation: Change of Stakeholder Requirements Model

In Hoorn (2004, Tech. Rep. [CD]), certain adjustments were made to the MoRC. While doing the first explorations in the field, I found that important business events seldom occur when you are around and that when they do,

managers do not want a nosy researcher in their way. It was hard to measure requirements for one or the other quality before and after an event took place because such events are unpredictable. Therefore, the research focus switched from events as the sole instigators of change to the goals and processes that were affected by these events. The (changed) goals and (changed) processes were supposed to regulate the level of agreement to requirements.

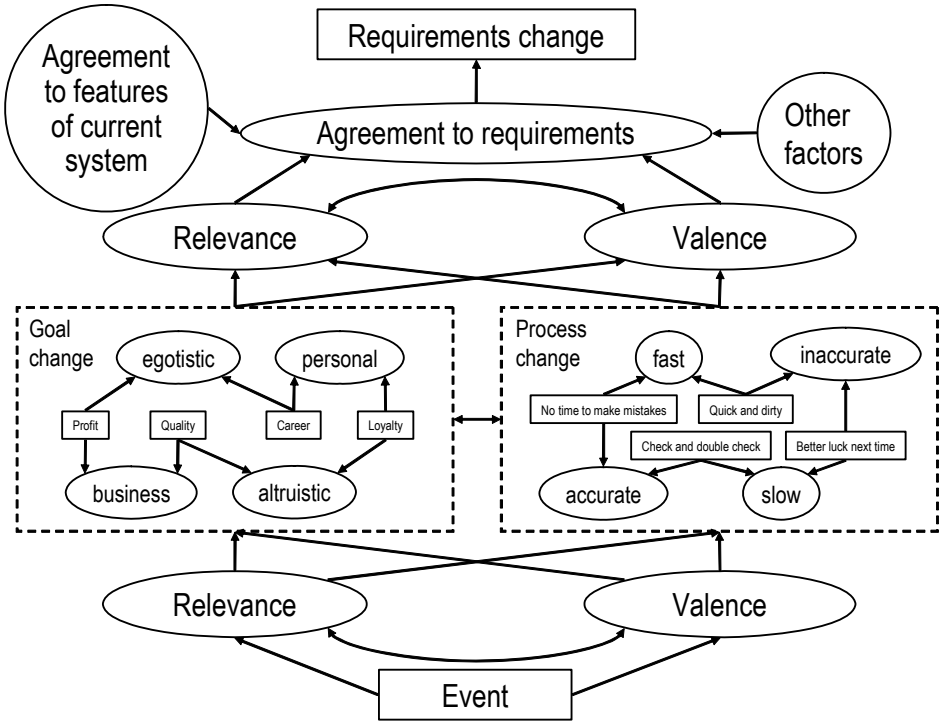


Figure 7: Change of Stakeholder Requirements model, an adaptation of the MoRC.

In so doing, however, assessment of the relevance and valence of events to goals became less important. Instead, assessment of the relevance and valence of requirements to goals and processes became germane. Therefore, another evaluative round was introduced to the model, in which the relevance and valence of requirements were estimated with regard to changed goals and changed processes. Finally, the factor 'agreement to the features of the current system' (cf. BTM1) was singled out from the container variable 'other factors,' being a serious candidate to affect the agreement to the requirements on the new system (BTM2). After all, working with the old system, Hoorn (2004, Tech. Rep. [CD]) supposed, will partially determine what people want from a new system. The renamed model Change of Stakeholder Requirements (CoStaR) is depicted in Figure 7.

The CoStaR model is somewhat more precise than the MoRC. It claims that through relevance and valence assessment, events may alter goals and processes (dashed boxes). More importantly, however, the requirements on a

system are again evaluated for relevance and valence in view of the changed goals and processes. Formulated as a chain of dependencies, events cause goals and processes to change through the mediation of relevance and valence to these goals. In addition, the changed goals and changed processes determine the level of agreement to requirements, again through the mediation of relevance and valence but now to the (possibly) changed goals. Agreement to the features of the current system may form an alternative or additional explanation of the variance in agreement to requirements. Therefore, agreement to the current system can be seen as a control variable.

2.11 Research in Businesses

In the chapters to come, I will verify many of the claims of the CoStaR model by conducting empirical research with real business cases. However, conducting research that is methodologically sound, yet readily applicable to real business problems is almost a contradiction in terms. At times, the differences between business and science place considerable stress on their relationship. Scientists blame businesses for staking their claims for product quality based upon unfounded and unjustified assumptions. Businesses blame scientists for doing research that is irrelevant outside the laboratory. In RE, such controversies are counterproductive. When problems arise in large information systems in large organizations, RE analysts will be forced to do science outside the laboratory. In the next sections, I analyze the differences between the business models of science and commerce, and conclude that the common interest lies in obtaining information about system and stakeholder requirements that is reliable as well as valid. I discuss in how far this can be achieved by applying controlled field experiments, combining a laboratory set-up with ecological validity. I will make a concise inventory of problems in scientific methods as a result of conducting experiments in businesses and how business receives scientific results. These issues are raised to appraise the results of the controlled field work in the chapters to come.

2.11.1 Different Business Models

Despite ongoing efforts of governmental agencies (e.g., in the Netherlands, Senter*Novem*'s IOPs)[4] and institutional discussion platforms (e.g., SIGCHI.NL),[5] there remains much to wish for in the collaboration between business and science in the information technology. It is not the lack of good intent, both parties see what is to gain from a well-oiled knowledge economy (e.g., integrating solutions, merging perspectives, and gaining commitment) (cf. Burrell, 2001). However, too many misunderstandings and wrong expectations keep the two worlds apart, which is a pity when leading to disempowerment and loss of innovation (cf. Burrell, 2001). The main point of concern, so

[4] http://www.senternovem.nl/iopmensmachineinteractie/index.asp
[5] http://www.sigchi.nl/

I argue, is that business and science work from fundamentally different business models, setting different goals and targets, which demand different types of business or work processes.

In a market economy, the business is in it for profit, which can be optimized by lowering operating costs, by better access to information, and quick response to events. University researchers work from a business model of continuation; they need to publish in order not to perish. The targets a business sets relate to selling a larger percentage of products than the last year, whereas the university's targets relate to producing as many scientific top publications as possible. Businesses do not care so much for the question why something works as long as it works. Universities and other research organizations do not only care for the how but also for the why. Without thoroughly tested explanations, they believe, there is no science and so they focus on laboratory work. Because knowledge institutions are so much preoccupied with the reliability and validity of information, their business processes are meticulously accurate (i.e. theoretical analysis, modeling, predicting, laboratory testing, evaluating) but at the cost of being relatively slow. To keep up with the competition, commercial enterprises cannot afford to cultivate slow business processes and quick fix solutions are often preferred over 'yet another round of painstaking check-ups.' Businesses are not so much concerned with reliability and validity of information but rather with issues of persuasion (e.g., advertisement, negotiation, and marketing). Businesses have very practical and acute matters to deal with in the here and now, demanding a hands-on (and therefore often non systematic) approach to problems that arise. Businesses often believe that scientists are solving toy problems because the scientists investigate but a limited number of factors for reasons of research-scope manageability and precision. Table 1 provides a summary of this analysis. For a deeper understanding of field versus laboratory work, consult Robson (2002, p. 12, Box 1.1 and p. 13, Box 1.2).

Table 1: Inventory of differences in business models of scientific and commercial organizations.

	Business	**Science**
Business model	Profit	Non-profit
Goals	Maximize profit	Continuation
Targets	Increase sales %	Increase publication %
Processes	Fast-inaccurate	Slow-accurate
Information should be	Persuasive	Reliable, valid
Problem solving style	Hands-on, non systematic	Lab-work, systematic control
Research focus	Know how	Know how and why

In constructing Table 1, I set out certain lines of supply and demand, discussed next. To maximize profit, it should be the business's concern to work with the most accurate information available. Universities are the most likely candidates for supplying such information, being reliability and validity prone. Part of the validity of information has to do with the question why a phenomenon occurs or what the underlying mechanism is. By concentrating on generic knowledge, scientific researchers are able to formulate one solution for many problems that seem quite different at surface level. On the one hand, the non systematic problem solving style of businesses could do with some systematic control as is usual in laboratory testing; lab-work, on the other hand, could leave its sterile, rigid, position to do some hands-on field experiments. Part of the profit a business makes, then, perhaps could be invested in the continuation of the research organizations (e.g., in the form of software, hardware, manpower, and providing a research field). With the more accurate information this generates, businesses can with more right center their persuasive activities on trustworthiness and credibility. Although research sponsorship should be seen as a long-term investment, yet, companies want to see the returns on their expenses within due time, and rightly so. This puts the burden on the universities to work faster than they are used to. However, increasing the speed of conducting research should not be at the cost of producing inaccurate results. If all goes well, the win/win situation, in the end, could be that in using the reliable and valid information that the universities provide, businesses can increase the number of sales events. The researchers can produce more and ecologically more valid publications with all that material from real-life cases. Money for quality information, so to speak, should be the deal between science and business. Ideally speaking. Ideally, because scientists may demonstrate repeatedly that a particular method seriously improves a software process or the performance of the organization. Yet, business managers prefer to follow the opinion of their professionals who supposedly are experts of the particular situation (Rainer, Hall, & Baddoo, 2003). Although management states to value empirical results, they favor the local opinion of the practitioners (ibid.).

2.11.2 Problems of Controlled Field Work

In the previous section, I have discussed some of the issues that separate businesses from knowledge institutions. Next I will make a small inventory of methodological problems that can occur when working in the field (science critique, cf. Robson, 2002, p. 8) and the applicability of the results from the business perspective (business critique, ibid.). These issues are raised to interpret and evaluate the controlled field work that is presented in the remaining chapters.

2.11.2.1 Science Critique

One of the main concerns in the scientific RE community is the issue of *small sample sizes*. For example, empirical software engineers tend to dispatch large numbers of different questions to several hundreds of stakeholders. Because of the large number of responses to one single item or question, these researchers believe they correct for response biases while gaining on generalizability of results.

Yet, such an approach may be possible in multinational companies and governmental institutions but not in medium or small size companies simply because no more personnel is available (sometimes fifteen people or less). The solution then is to downsize the scope of the research and to ask many of the same type of questions (increasing the number of observations) about a limited number of topics. This, moreover, improves the reliability of the results because the researcher can account for variability in the responses within stakeholders. This is not possible with single-item responses in spite of larger sample sizes. Admittedly, with a small sample you spend most of your time trying to find a stable and viable result. Yet, if you can establish strong significant effects in a small group of people, the effects must be robust. In that case, generalizability of results is unproblematic because the power of effects will only gain from a larger sample size. Thus, customary engineers gain less reliable knowledge over more issues whereas I argue for gaining more reliable knowledge over fewer issues.

Another point is that RE analysts are used to carry out interviews or protocol studies on the stakeholders' decision making process. This is to interpret the requirements or changes in the requirements that they observe at the work floor. Without such information, they believe, research results are conjectures at best or artifacts of experimental construct which will bias the result towards your derived models. This is the issue of *lack of control of the meaning of items*. Of course, an interview may sketch some background information on why and how requirements were elicited or changed (cf. Robson, 2002, p.). There is nothing wrong with that, however, in leaving it there, the bias is not in the experiment but in the interviewer who (not deliberately) guides the stakeholder to or from earlier derived models (cf. Shaw, 2002; Robson, 2002, pp. 456-473). In addition, interviews and to a lesser extent protocol studies hardly have rigorous means to control, let alone, improve the reliability of their measures (ibid.). Again, by repeating observations within stakeholders, reliability of the measures can be calculated (e.g., Chapter 4) or even improved by repairing a poorly performing item (or question) or deleting it from analysis. Yet, businesses do not always allow you to pretest or retest the items on your questionnaire (too much bother for the personnel). In that case, controls have to be performed post hoc.

Controlling in how far items measure one and the same concept and not another concept also excludes the necessity to know the exact relationships between a specific requirement and a specific goal as well as the need to work

from an exhaustive requirements list. The aim of acquiring general RE knowledge is not to engineer one specific system for one specific company. A small list of requirements and goals suffices as long as they reliably measure the relevant variables, such as the musts on a system or the goals stakeholders want to achieve with it. The type of relations between the *general* variables and their strength (e.g., between must requirements and goals to achieve) are established statistically, describing how this relation should be conceived of in subsequent systems and stakeholders.

The third issue is *the need for replication* to control for spurious findings (e.g., Ohlsson & Runeson, 2002). This is one of the hardest parts in empirical RE. It is not at all obvious that a company will allow you to repeat the research and to replicate results (the project's timeline is sacred). Companies believe that you already took more than enough time of their employees to do your work. They think you are a bad researcher if you did not do a proper job the first time. Or something can happen between the pilot study and the main study, as in Chapter 3, where the business abdicated the system entirely so that the complete case was gone. To save the main study, the only thing to do was to conduct RE for a to-be-developed system instead of a system-under-development.

This illustrates that in the field, conditions hardly can be maintained stable across replication studies. If you are forced to repeat results in other settings, in other businesses, with different types of stakeholders, results may easily alter, deteriorate, or die. In Chapter 4, for instance, the requirements in the first case study did not fit the views on the stakeholders, so that one could argue that only negative associations were observed, which do not allow for proposing a general model from this instance. Conversely, if you are capable of repeating results throughout changing research conditions, you must have one hell of a point (i.e. the other cases in Chapter 4 and one in Chapter 5).

2.11.2.2 Business Critique

While the industry will only have a few complaints about the experimental methods and statistical techniques that you apply, they have great difficulty in seeing how those methods and statistics could benefit the industrial (RE) community. For the industrial audience the case is simple: The proposed methods are usually too dense, too difficult to understand, and too labor intensive to ever get used. But then again, they don't need to. This problem of the *business utility of scientific methods* is based on a false impression. It is not the task of businesses to do such research; that is what the academy is for. It is the academic task to provide solid research results, which can later on be translated into focal points of RE in practice, possibly apprehended with more lightweight approaches.

Related to the utility problem is that well established methods in the RE literature such as VORD (Viewpoint Oriented Requirements Definition) (Kotonya & Sommerville, 1992), KAOS GRAIL (Bertrand et al., 1998), Sce-

nIC (Potts, 1999), and in industrial practice Volere (Robertson & Robertson, 2004) could all do the job just as well as statistically intensive techniques. The industry wants to know how to justify the extra 'bangs for bucks' that a statistically motivated approach delivers. In other words, the industry fears a *lack of added value*. Well, it is not a matter of producing elegant statistics to outdo existing RE methods. The good news is exactly that Volere etc. can establish the same results. Point is, however, that they never did and that the statistical approach advanced in this book uncovered relationships and possibilities never contemplated before. Again, it is the academic task to develop a reliable, valid, and general theory of requirements (here, a theory of requirements change), which may guide the follow-up in the industry by means of statistically less intensive approaches such as Volere or VORD.

2.11.3 Conclusions/Discussion

I have attempted to create something of an eye-opener to certain blind spots in the relationships between commerce and science. Yes, businesses want to make money and scientists want to publish papers but both need ecologically valid and methodically reliable data to increase their performance.

Reliability and validity of the data, then, should be a treasure safeguarded by both parties. Scientists need such data to contribute to high-ranking journals; businesses need them to adapt their products and to focus their persuasion activities on trustworthiness and credibility. In the area of RE and system design, controlled field experiments seem to provide the common ground for research and development (R&D) in both science and business. Such experiments allow for systematic observations under the strictest conditions that are possible in the field more than hands-on experience can offer. This can be done, moreover, while preserving higher ecological validity than any lab-work can warrant.

With regard to the critique of both the academic and business RE community on the field-experiment approach and the use of heavyweight statistics, I would like to comment in the following way. I believe that it is important to establish a sound theoretical basis for certain issues in RE – if necessary with elaborative and resource intensive methods. Then these results could guide RE in practice while employing the more lightweight approaches such as Volere (Robertson & Robertson, 2004). The reasons why these methods and statistical techniques should be used are to improve the reliability of the observations and to allow for generalization across systems and stakeholders. That way, we do not have a series of more-or-less incomparable case studies to derive information from but a series of systematic tests that provide RE knowledge in a controlled and repeatable way (Shaw, 2002). According to Rainer, Hall, and Baddoo (2003), practitioners and business people prefer local opinion, then empirical evidence from the business case at hand, and finally external empirical evidence. This makes you think of how the present thesis will be received

and urges to make sure that general findings can be translated back into specific business cases while sustaining or countering common opinion.
Let me end with some warnings and suggestions on how to improve research collaboration between business and science. These comments are based on real-life field experience and explicitly making dirty hands (cf. Robson, 2002, p. 508).

On the side of the business, give scientists the opportunity to employ their thoroughness. Business secrets and hushing up problems may be understandable but do not supply reliable information. If possible, try to sell your application with the possibility for the client to have their newly purchased system scrutinized by scientists for free. This has worked, for example, in establishing a contract between The Mediator Group and IBM (Chapter 3). Be generous in supplying software, hardware, possible research assistants, programmers, and other infrastructure. Moreover, try to actively involve possible participants or respondents. Because scientists work with relatively low budgets, supply of equipment and other resources increases quality and speed of the research. In this context, VirTouch supplied enough VTPlayer mice, which are extremely expensive, to arrange parallel experimental sessions with a sample of blind pupils (Chapter 4). This improves research speed and reduces the chance that stakeholders will influence each other by talking about the experiments. The returns are that businesses get top R&D for relatively low costs and that they are in the front line of scientific innovation and thus have a pre-competitive advance to other players. However, do not have too high an expectation of the results. Experiments can fail and results may be disadvantageous to marketing strategies. Sometimes, systems do not yield the expected effects or even deteriorate an existing situation. Scientists should be able to make such information known, nevertheless.

On the side of science, reckon with the fact that the field is ever changing. For example, The Mediator Group (Chapter 3) started their contract with a financial sub department of IBM, which then became an independent business. Account for such events and represent changes in business models as factors in scientific models that are dynamic, that is, describing situations that probably never reach a stable end state. Be careful with highly sensitive information and situations. Do not get in the way of negotiations between business and client and do not violate business agreements or workflow for the sake of scientific rigor. Work efficiently. Try to be as invisible as possible on the work floor by doing a lot of ground work at your office or behind the screen and then perform a hit-and-run experiment. This least interferes with the business's normal way of working and minimizes the bias of experimenter effects. Stay in your role. A scientist should be a relatively neutral observer and not become part of a design team or act as an external consultant to the management board.

In this book, my concern is to satisfy both business and science by "turning research questions into projects" (Robson, 2002, p. 79). The aim during my studies was to find a "win/win" solution, which I found in administering

controlled field experiments. By learning from each other and improving our relationships, business and science can support "open discussion of issues, task proficiency, equal distribution of work amongst the team members, better brainstorming, and development of creative problem solving. This asks for the ability to use active or effective listening, confront situations in a non-threatening way, analyze input, and identify underlying concerns" (Burrell, 2001).

2.12 The Remainder of this Book

The focus of the empirical work in the coming chapters is not so much on the events that alter goals and processes but right after, when goals and processes have changed and requirements are re-evaluated. First, this was a purely practical matter. As said, the companies and organizations that I worked with did not allow me to sniff around and ask all kinds of questions while they were in the middle of a buy out operation or reorganization. Second, the main question of my assignment for the agency of Economic Affairs, Senter*Novem*, was how changes in (business) goals and processes could lead to requirements change. Having said this, let's have hands-on experience with lab-work in the IT field and proceed with Chapter 3 on relevance of goals and requirements prioritization, Chapter 4 on agreement with requirements, and Chapter 5 on stakeholder satisfaction, usability, and efficiency requirements.

References

Arnold, M. (1960). *Emotion and Personality.* New York, NY: Columbia University.

Bertrand, P., Darimont, R., Delor, E., Massonet, P., & Van Lamsweerde, A. (1998). GRAIL/KAOS: An environment for goal driven requirements engineering. *Proceedings of ICSE'98 - 20th International Conference on Software Engineering*. Kyoto, Japan: IEEE-ACM.

Bhulai, A., & Van der Veer, G. C. (2005). *After All, the Web is Just like the Real World.* Submitted.

Burrell, B. (2001). *Conflict Management.* Retrieved April 3, 2004 from http://web.mit.edu/ collaboration/mainsite/modules/module1/1.11.5.html

Chisăliță, C., Puerta Melguizo, M. C., Hoorn, J. F., Van der Veer, G. C., & Kok, E. (2005). Cultural differences in user groups: A multi-angle understanding of IT use in large organizations. *Cognition, Technology, & Work, 7,* 101-110.

Cooper, A., & Reimann, R. (2003). *About Face 2.0: The Essentials of Interaction Design.* New York, NY: John Wiley & Sons.

Darimont, R., & Van Lamsweerde, A. (1996). Formal refinement patterns for goal-driven requirements elaboration. In D. Garlan (Ed.), *Proceedings of the Fourth ACM SIGSOFT Symposium on Foundations of Software Engineering* (pp. 179-190), San Francisco, California, USA, October 16-18, 1996. ACM SIGSOFT Software Engineering Notes 21(6), November 1996).

De Sousa, R. (1987). *The Rationality of Emotion.* Cambridge, London: MIT.

Döring, B., Dörfel, G., & Distelmaier, H. (2001). Acquisition and modeling of a priori knowledge for decision support systems. In Onken, R. (Ed.), *CSAPC '01, 8th*

Conference on Cognitive Science Approaches to Process Control (pp. 211-220). Bonn, Germany: Lilienthal-Oberth.

Frijda, N. H. (1986). *The Emotions.* New York: Cambridge University.

Hirschheim, R., & Klein, H. K. (1989). Four paradigms of information systems development. *Communications of the ACM, 32, 10,* 1199-1216.

Hoorn, J. F. (2004). *Rationale for REquest, a Structured Questionnaire for General Requirements Validation.* (Tech. Rep.) [CD]. Amsterdam: Free University.

Hoorn, J. F., Van Dam, S., Fambach, G., Van Nieuwland, A., & Van der Veer, G. C. (2004). Looking for win/win solutions between lab-work and hands-on experience in IT research. *Proceedings of SIGCHI.NL 2004: Dutch Directions in HCI* [CD]. New York: ACM Press.

Hoorn, J. F., & Van der Veer, G. C. (2003a). Requirements analysis and task design in dynamic environments. In D. Harris, V. Duffy, M. Smith, & C. Stephanidis (Eds.), *Human-centred computing: Cognitive, social, and ergonomic aspects. Volume 3 of the proceedings of HCI International 2003,* Hersonissos, Crete, Greece (pp. 472-476). Mahwah, NJ: Erlbaum.

Hoorn, J. F., and Van der Veer, G. C. (2003b). Requirements engineering for groupware to manage business process dynamics and stakeholders' mindset. In G. C. van der Veer, & J. F. Hoorn (Eds.), *Proceedings of the 9th European Conference on Cognitive Science Approaches to Process Control: Cognition and Collaboration - Distributed Cognition in Complex Processes 2003,* Amsterdam, Netherlands (pp. 153-160).

Konijn, E. A. (2000). *Acting Emotions.* Amsterdam: Amsterdam University.

Kotonya, G., & Sommerville, I. (1992). Viewpoints for requirements definition. *Software Engineering Journal, 7,* 6, 375-387.

Kotonya, G., & Sommerville, I. (1998). *Requirements Engineering: Processes and Techniques.* New York, NY: John Wiley.

Laurel, B. (1991). *Computers as Theatre.* Boston, MA: Addison-Wesley.

Lazarus, R. S. (1982). Thoughts on the relations between emotion and cognition. *American Psychologist, 37,* 1019-1024.

Lundberg, M., Goldkuhl, G., & Nissen, N. (1981). *Information Systems Development: A Systematic Approach.* Englewood Cliffs: Prentice-Hall.

Natt och Dag, J., Gervasi, V., Brinkkemper, S., & Regnell, B. (2005). A linguistic-engineering approach to large-scale requirements management. *IEEE Software, 22, 1,* 32-39.

Ohlsson, M. C., & Runeson, P. (2002). Experience from replicating empirical studies on prediction models. *Proceedings of the 8th IEEE International Symposium on Software Metrics (METRICS'02),* Ottawa, Canada (pp. 217-226). Los Alamitos, CA: IEEE Computer Society.

Potts, C. (1999). ScenIC: A strategy for inquiry-driven requirements determination. *Proceedings of RE'99: International Symposium on Requirements Engineering.* Limerick, Ireland.

Price, J. and Cybulski, J. (2004). Influence of stakeholder communication on consensus making in requirements negotiation. In *AWRE '04 Conference Proceedings of the 9th Australian Workshop on Requirements Engineering.* Adelaide, Australia: University of South Australia.

Rainer, A., Hall, T., & Baddoo, N. (2003). Persuading developers to 'buy into' software process improvement: Local opinion and empirical evidence. *Proceedings of*

the International Symposium on Empirical Software Engineering (ISESE'03) (pp. 326-335), September 30 – October 1, 2003. Rome, Italy.

Rauschenbach, U., & Schumann, H. (1997). *Adaptive Image Transmission.* WSCG '97 Winter School of Computer Graphics (February 10-14), Plzen, Czech Republic.

Rigdon, E. E. (1998). Structural equation modeling. In G. Marcoulides (Ed.), *Modern Methods for Business Research* (pp. 251-294). Mahwah, NJ: Lawrence Erlbaum.

Robertson, J. & Robertson, S. (2004). VOLERE: Requirements specification template. Edition 10.1. Atlantic Systems Guild.

Robson, C. (2002). *Real World Research. A Resource for Social Scientists and Practitioner-Researchers.* Malden, Mass.: Blackwell.

Sartre, J. P. (1936). Esquisse d'une Théorie des Émotions. Paris, France: Hermann.

Sebillotte, S. (1995). Methodology guide to task analysis with the goal of extracting relevant characteristics for human-computer interfaces. *International Journal of Human-Computer Interaction, 7, 4,* 341-363.

Shaw, M. (2002). What makes good research in Software Engineering? *International Journal of Software Tools for Technology Transfer, 4, 1,* 1-7.

Singer, T. (1996). Stop thief! Are your employees robbing you blind? *Entrepreneur Magazine, 1,* 148-153.

Steensma, H. (1999). Psychologen over rechtvaardigheid. Theorie, empirie, en nieuwe ontwikkelingen. *BandM, Tijdschrift voor Beleid, Politiek en Maatschappij, 26, 2,* 106-121.

Tauber, M. (1988). On mental models and the user interface. In G. C. van der Veer, T. R. G. Green, J.-M. Hoc, & D. Murray (Eds.), *Working with Computers: Theory versus Outcome* (pp. 89-119). London, UK: Academic.

Usability Glossary (2005). *Speed-accuracy Tradeoff.* Retrieved March 09, 2005 from http://www.usabilityfirst.com/ glossary/term_475.txl

Van Leeuwen, E. (2001). Preserving Identity when Groups Combine. A Study of Group-based Reactions to Mergers. Leiden, NL: Leiden University.

Van der Veer, G. C., Lenting, B., & Bergevoet, B. (1996). GTA: Groupware Task Analysis – Modeling Complexity. *Acta Psychologica, 91, 3,* 297-322.

Van der Veer, G. C., & Van Vliet, H. (2001). The human-computer interface is the system; a plea for a poor man's HCI component in software engineering curricula. In D. Ramsey, P. Bourque, & R. Dupuis (Eds.), *Proceedings 14th Conference on Software Engineering Education & Training, IEEE* (pp. 276-286). Piscataway, NJ.

Van Welie, M. (2001). *Task-based User Interface Design.* Amsterdam: Free University.

Van Welie, M., & Van der Veer, G. C. (2000). Structured methods and creativity – a happy Dutch marriage. In *Proceedings of Co-Designing* (pp. 111-118). Coventry, UK.

Wieringa, R. J. (1996). *Requirements Engineering: Frameworks for Understanding.* New York, NY: John Wiley.

Zlatev, Z., Van Eck, P., Wieringa, R. J., & Gordijn, J. (2004). Goal-oriented RE for e-services. In L. Baresi, M. Papazoglou, & R. Wieringa (Eds.), *Proceedings of the International Workshop on Service-oriented Requirements Engineering (RE'04).* Kyoto, Japan: IEEE CS.

3 Relevance and Prioritization

Abstract

Stakeholders judge goal relevance and software development-project requirements differently considering them from a business versus a personal viewpoint.[1] Three empirical studies provided evidence that stakeholders' personal goals for a system are valued higher than business goals and that, moreover, changes in personal goals lead to significantly more prioritization activity than changes in the business goals. In Study 1, police officers expressed their agreement to a list of requirements on a Capacity Management System from a business and a personal viewpoint. I found a 'requirements-analysis rift,' indicating that stakeholders regard requirements as business-related, agreeing with them even if these requirements do not match their personal goals. In Study 2, science students rank-ordered the same list of requirements on an e-learning environment from a university and a personal perspective. In Study 3, a similar study was conducted with a sample of financial consultants. Both studies revealed that, when personal goals changed, prioritization shifts were stronger than when business goals changed. The factor of egotistic-altruistic goals rendered merely insignificant effects and was replaced by the more effective factor 'Model Change.' Further, I suggest a method using the Spearman rho statistic to calculate changes in requirements prioritization under different conditions of goal change.

Keywords: Requirements, Priorities, Change, Viewpoints, Business, Goals, Measurement

3.1 Introduction

People change their minds all the time. However, when they change their minds about a large computer-based system under development, trouble may ensue. The simple question of adding a 'multiple undo' feature to a graphics editor impacts the structure of that system deeply, entailing more work, more negotiation, more time, and eventually, greater costs. In addition, the later in the development cycle change requirements emerge, the more costly the system's revision becomes (Boehm & Pappacio, 1988).

The literature frequently sums up a number of factors that cause requirements change (e.g., Lehman, 1996). These factors may be social (e.g., imper-

[1] This chapter is based on Hoorn and Breuker (2005) and Hoorn, Breuker, and Kok (2006).

fect communication), technical (e.g., new developments), or organizational (e.g., business model change). Additionally, an abundant literature proposes methods to describe (e.g., Potts & Takahashi 1993), analyze (e.g., Gervasi & Zowghi 2005), and handle requirements change (e.g., risk assessment in Strens & Sugden 1996). To my knowledge, however, no "key hypotheses" exist to guide such efforts beyond instructions on *how* to detect volatile requirements and prevent a system from becoming outdated. In this chapter, I formulate five hypotheses that address some of the reasons *why* requirements change and how applicable factors affect changes in requirements prioritization. The remainder of this introduction develops the hypotheses based on issues the literature raises.

Stakeholders involved in software development change their minds for many reasons. The occurrence of a market event that changes the business model, which affects the system requirements, is often mentioned as a cause (Highsmith & Cockburn, 2001; Cockburn & Highsmith, 2001). A software production-house may become a service-oriented organization, changing the entire IT infrastructure. A recession could be such an event, leading a business from hardcore commercial activities to more altruistic business goals (BISON, 2003). A business-model change can alter requirements prioritization to align the supporting IT with the new goals (Lam et al., 1999). Thus, certain requirements lose their relevance. In view of a new situation, requirements that were important may no longer be, whereas others may move to the top of the list. With respect to personal goals (cf. Loucopoulos & Kavakli, 1995), people may shift from more selfish goals (e.g., shielding off information, workload reduction) to more altruistic ones (e.g., sharing files, collaborative work) (cf. BISON, 2003; Hopkins, 2002). That is, not only do business models change; 'personal models' may change as well.

Due to the many shifts in foci and priorities, stakeholders sometimes seem inconsistent about what they actually want to accomplish with the system (Van Lamsweerde, 2004). Discrepancies can occur between business and personal models. Cooper and Reimann (2003) indicate that interactive systems are more successful when designed not only considering the business point of view but users' personal goals as well. Price and Cybulski (2004) state that stakeholders "often value their personal goals over those of their employer," making it important to align requirements with personal goals to prevent software development-project failure. In other words, agreement to requirements may depend on the viewpoint taken (business or personal).

Study 1 explores how viewpoints (i.e. business vs. personal) can be used for requirements validation. I utilize several viewpoints approach concepts to study how the relevance or irrelevance of goals relates to agreement with requirements. The viewpoints approach to requirements engineering tries to integrate more perspectives on requirements into the system specification (e.g., Easterbrook & Nuseibeh, 1996). "Viewpoints are entities which may be used to structure the process of requirements elicitation and to structure the re-

quirements specification" (Sommerville & Sawyer, 1997). It does so to detect and resolve possible inconsistencies at an early stage of software development.

The first viewpoints hypothesis (H1) posits that personal goals to achieve with a system are regarded as more urgent, grave, or relevant (Frijda 1986, p. 494, p. 463) than business goals. Moreover, (H2) requirements insufficiently aligned with personal needs evoke less agreement from a personal than from a business point of view.

In Studies 2 and 3, I investigated the effects of different viewpoints on requirements change, in particular, the changes in how requirements are prioritized. These prioritization studies surveyed three additional hypotheses.

H3 states that changes in goals (business or personal) will change the relevance of requirements, reflected in changing the priorities on the requirements list. In line with BISON (2003) and Hopkins (2002), H4 suggests that changes in requirements priorities are sensitive to changes from selfish to altruistic goals and v.v. H5 is an alternative to H4, expressing that the transition from old to new goals affects the relevance (i.e. priorities) of the requirements rather than a change in the type of goals.

The remainder of this chapter is structured as follows. In Study 1 (Section 3.2), police officers indicated their agreement with a list of goals and requirements. Goals could be relevant or irrelevant within a given viewpoint (business or personal). In Study 2 and 3 (Sections 3.3 and 3.4), prospective users of an e-learning system prioritized a list of requirements regarding the business goals or their personal goals. The general conclusions and discussion in Section 3.5 are that stakeholders do not necessarily connect requirements to personal goals, but changes in personal goals impact requirements prioritization more strongly than changes in business goals do. I consider the possible lessons learned from the empirical 'viewpoints research' for requirements engineering in practice.

3.2 Study 1: Goal Relevance and Agreement to Requirements

These studies were directed at viewpoints related to system stakeholders (Sommerville & Sawyer, 1997). "A viewpoint is an encapsulation of partial information about a system's requirements" (ibid.). It is derived from a standing or mental position by an individual (Leite & Freeman, 1991) or group of individuals.

Together with Evelien Kok, at that time working at the Concern Information Management Police (CIP), I investigated the police management perspective on the requirements for a Capacity Management System (CMS), used for scheduling tasks and allocating personnel. I also studied the perspective of the prospective end-users of that system, the police officers. Moreover, we obtained the *viewpoint concerns* (ibid.), consisting of the organizational goals, business targets, and limitations of the management as well as the personal goals and restrictions of the officers. For both business and personal goals, we collected goals with a more altruistic or a more egotistic or selfish flavor (cf.

BISON, 2003; Hopkins, 2002). The *viewpoint requirements* (Sommerville & Sawyer, 1997) were kept constant throughout viewpoints. These were the requirements the CMS 'must have' and 'won't have' as negotiated and agreed-upon by the management to achieve strategic business goals.[2]

We wanted to see how much future end-users of the CMS agreed with management-viewpoint requirements (must vs. won't) from a personal or a business standpoint. Police management mainly provided CMS requirements meant to sustain a new strategic mission. We hypothesized (H1) that the police officers would agree with personal goals more than the business goals (Price & Cybulski, 2004) and thus (H2) disagree with the management-viewpoint requirements more from the personal perspective than from a business viewpoint.

3.2.1 *Method*

3.2.1.1 Participants and Experimental Design

Students of the Dutch Police Academy (N= 33; 22 male, 11 female; age 19-45, M= 26.5, SD= 5.96; years in service M= 2, SD= 1) participated in a questionnaire study that concerned the redesign of a Capacity Management System for allocating workforce to a task, planning actions, and scheduling holidays and shifts. These participants ranged from the same district and functions within the organization (officer or chief officer). They were already working but studied at the academy one day a week and had some experience with the CMS. In an adaptation of the stakeholders analysis template offered by Alexander and Robertson (2004), Table 1 provides stakeholders' profiles of the officers and their managers in terms of their roles, responsibilities, success criteria, and involvement with the CMS.

[2] The MuSCoW list speaks of requirements that Must be, Should be, Could be, or Won't be on the system (eRA, 2002).

Table 1: Stakeholders' profile of officers and managers.

Stake-holders	Descrip-tion	Type	Respons-ibilities	Success criteria	Involve-ment	Deliv-erables	Issues
Officers	On duty in the streets, adminis-ter activi-ties be-hind the desk	Know how to deal with real life incidents, low on abstract problem solving	Help people in the streets, solve crimes, to serve and protect	Number of crimes solved, number of fines written	High with policing, low with adminis-trative tasks	Day reports, hour ad-ministra-tion, incident adminis-tration	Want to be more on the streets and not behind the desk
Managers	Plan and co-ordinate actions, control officers	Hierarch-ical posi-tion, think in terms of factors rather than people	Managing the pre-cinct, keeping in step with politics, solving societal urgent matters	Keeping the per-formance contract with the govern-ment	High with adminis-tration and fi-nance, low with the offi-cers well-being	Month and year reports, financial reports, overall perform-ance evalu-ations	Want to have more control over the time and effort spent on a case

With respect to Table 1, the CMS was created as a management tool for implementing established rules. The management did not believe officers' goals and requirements to be significant for scheduling purposes. Yet, the officers were key stakeholders in the CMS because of their role in filing system data and the potential impact of scheduling and planning on their work (e.g., in showing initiative, holiday privileges, or time behind the desk). Therefore, we were interested in future end-users' agreement with requirements about which they had not been consulted.

Four between-subjects conditions were established, each approaching the same list of management-viewpoint requirements from a different perspective. The officers indicated a level of agreement with goals within a viewpoint, which could be Business or Personal (factor Stakeholders' View). These goals could be Egotistic or Altruistic (factor Goal Orientation). The within-subjects factors were Goal Relevance (Relevant vs. Irrelevant) and Requirements (Must vs. Won't).

3.2.1.2 System

As is, officers of the Dutch police force have to justify the hours they work during their shifts (Table 1). Each organizational police unit has planners who schedule the personnel. This rather complex task involves several different shifts (morning, evening, and nightshift, as well as weekend and stand-by shifts). The law designates how many hours and nightshifts an officer is allowed to work within a certain timeframe. Police officers often work extra hours, for example, when incidents occur or when colleagues are absent, re-

sulting in even more complex planning. Several systems that support planning and schedules in the police organization exist. In the early nineties of the last century, a majority of the police districts implemented a tool called Registration Planning and Control (RPC) to support planning and registration. To date, the planner registers schedules, officers update/confirm, and their chief authorizes them, all through RPC. RPC is linked to the salary system, so that the registered working hours are directly related to the salary. Other precincts, however, have no central system for planning and control.

In the district we investigated, the planner makes schedules in a stand-alone system and officers register their hours in a spreadsheet. The follow-up of the RPC is the CMS basic tool, recently implemented and now used by some of the officers. The officers consider hour registration a necessary evil—as it relates directly to salary, officers are motivated to use it. The old RPC system is character-based and user-unfriendly, whereas stand-alone systems are inefficient and error prone. One mistake can lead to incorrect or even unpaid salaries. The CMS was built to improve on these matters.

3.2.1.3 Procedure

Working with Evelien Kok as functional analyst of the CIP, in-depth ethnography established a list of requirements for the CMS desired by the corps management. Moreover, we also acquired a list of these managers' business goals as well as a list of personal goals of the officers (not the same people who participated in the questionnaire study). Business and personal goals were categorized as either egotistic/selfish (e.g., "I want to have freedom") or altruistic (e.g., "We want to serve society") (cf. Table 1). Categorization was based on personal interviews and participatory observations and this information was used in a structured Requirements Engineering questionnaire, the CMS RE*quest* (Appendix 3.1). The questionnaire, written in Dutch, included 79 items and two open-ended questions. After receiving an introduction, officers filled out the paper-and-pencil CMS RE*quest*, which was divided into 4 blocks. The first block consisted of 24 items on the goals and concerns of the officers or the managers. The second block of 24 items concerned the requirements for the new CMS. The third block of 24 items systematically connected a goal to a requirement while attaching a valence. The results of this measurement are reported in the next chapter (Section 4.3). The fourth block consisted of 7 socio-demographic items (age, function, etc.) followed by two open-ended questions.

The police officers were randomly assigned to one of four versions of the CMS RE*quest* (Appendix 3.1). In the P_e version (n= 8), the CMS RE*quest* focused on Personal goals with an Egotistic thrust (e.g., "I want to keep my holiday privileges"). In the P_a version (n= 8), the Personal goals were more Altruistic (e.g., "I want to support my colleagues"). The B_e version (n= 9) focused on Business goals that only served the organization's Egotistic aspirations (e.g., continuity, cost-effectiveness), whereas in the B_a version (n= 8),

the Business goals were Altruistically oriented (e.g., to serve and protect society). The management-viewpoint requirements, however, were the same in all four versions. Thus, we could examine how each viewpoint, manifested in the different types of goals, affected officers' agreement with requirements.

Note that all items were presented to the officers as affirmative statements to avoid answering biases and response confusion that could result from using linguistic negations (Dillman, 1999). That is, won't requirements and irrelevant goals were phrased desirably, and the officers were expected to disagree to the won't (put as must) requirements and the irrelevant (put as relevant) goals. Items were pseudo-randomly distributed within blocks.

In a classroom setting, the police officers received an introduction and were asked to fill out the questionnaire. This was done for two separate classes—a second year and a third/fourth (= final) year class. Completing the questionnaire took between 15 and 20 minutes.

3.2.1.4 Measurements

We developed eight unipolar relevance scales, keying goals that supposedly were relevant or irrelevant to the work of the officers as derived from the ethnographic study. These goals could be Personal Egotistic (P_e), Personal Altruistic (P_a), Business Egotistic (B_e), or Business Altruistic (B_a). The Likert items on these scales were scored for agreement on a 6-point rating scale (0= completely disagree, 5= completely agree).

Items that featured personal goals were introduced by a line that said "I find it important that …" followed by a possible completion, for instance, "I can work in a team." Below is a bulleted list of sample items.

P_e Relevant scale, 12 items
I find it important that…
- I keep my holiday privileges
- I have freedom to show initiative

P_e Irrelevant scale, 12 items
I find it important that…
- privileges are ignored
- initiative is discouraged

P_a Relevant scale, 12 items
I find it important that…
- I consider the wishes of my colleagues
- I have the time to help my colleagues

P_a Irrelevant scale, 12 items
I find it important that…

- I can ignore the wishes of my colleagues
- colleagues solve their own problems

B_e Relevant scale, 12 items
I find it important that my corps…
- makes a professional impression on the outside world
- saves money on allocating personnel

B_e Irrelevant scale, 12 items
I find it important that my corps …
- makes an amateurish impression
- spends more money on personnel

B_a Relevant scale, 12 items
I find it important that my corps …
- decreases distance towards civilians
- fights terrorism

B_a Irrelevant scale, 12 items
I find it important that my corps …
- increases distance towards civilians
- leaves the war on terrorism to the national government

We also created two unipolar scales to measure agreement with the management-viewpoint requirements on a 6-point rating scale.

Requirements Must scale, 12 items
- Schedules are definite 48 hours in advance
- Schedules are arranged on the basis of expected activity

Requirements Won't scale, 12 items
- Schedules can change continuously
- Schedules are arranged on the basis of available personnel

The items on all these scales were tested by focus groups for readability, wording, and whether their contents made sense to people working in the field. After necessary changes had been made, a focus group again inspected items, after which we considered them ready for the main test.

3.2.2 *Analysis and Results*

3.2.2.1 Scale Analysis

Each scale originally consisted of 12 items. To improve the measure's reliability, I selected items based on Corrected Item-Total Correlations and standard-

ized Cronbach's alpha. Items correlating < .01 or negatively with the scale total were removed until alpha > .60. Moreover, the standard deviation of each item should be around 1, and skewness of items and scale < .60 (for details, see Hoorn & Kok, 2005, Tech. Rep. [CD]). I wanted to establish at least two items on a scale that did not correlate strongly with other scales (discriminant validity). Thus, I calculated the total scores of each scale and conducted Pearson correlations between each item with the total scores of the other scales. I compared the Corrected Item-Total Correlations of each scale item with the Pearson correlations to select those items that correlated stronger with their own scale than all other scales. I then reanalyzed the reliability of the final scales (results are tabulated in Table 2).

Scale length ranged from two to four items, and Cronbach's alpha of all shortened scales was acceptable (.67) to excellent (.98) (Table 2). Discriminant validity of the scales was also excellent in more than 75% of the cases. Discriminant validity was modest in two cases (P_e Relevant scale, B_a Requirements Won't scale) and poor in one case (P_a Irrelevant scale), indicating considerable correlation with other scales and therefore rather indistinct measurements.

Table 2: Standardized Cronbach's α, discriminant validity (d), and number of items (#) on a scale for four questionnaire versions (N= 33).

Questionnaire version	**Personal Egotistic n= 8**			**Personal Altruistic n= 8**			**Business Egotistic n= 9**			**Business Altruistic n= 8**		
Metric	**α**	**d**	**#**	**α**	**d**	**#**	**α**	**d**	**#**	**α**	**d**	**#**
Relevant scale	.71	±	2	.86	+	3	.84	+	4	.83	+	3
Irrelevant scale	.83	+	3	.78	–	2	.77	+	2	.77	+	3
Requirements Must	.98	+	3	.67	+	2	.74	+	2	.67	+	2
Requirements Won't	.86	+	3	.81	+	4	.78	+	3	.80	±	2

Note. + = all items on the scale show good discriminant validity, ± = one item shows poor discriminant validity, – = all items show poor discriminant validity

3.2.2.2 Agreement to Goal Relevance and Requirements

I analyzed the effects of different viewpoints on the level of agreement with the relevance of goals and the level of agreement with the requirements that should satisfy those goals. The four different versions of the CMS RE*quest* represented a 2 (Stakeholders' View: Business vs. Personal) * 2 (Goal Orientation: Egotistic vs. Altruistic) between-subjects design. The within-subjects factors were Goal Relevance (Relevant vs. Irrelevant) and Requirements (Must vs. Won't). However, it could be argued that answering biases occurred in responses to indicative items (relevant goals, must requirements) compared to contra-indicative items (irrelevant goals, won't requirements). In addition, the Goal Relevance and Requirements scales could not be considered inde-

pendent measurements (sometimes the discriminant validity was modest – Table 2).

Therefore, I devised a factor called Item Type that contrasted the indicative items (relevant and must) with the contra-indicative items (irrelevant and won't). Scales was another overall factor, which contrasted the whole of Goal Relevance with the whole of Requirements. This way, the interaction between Item Type (Indicative vs. Contra-indicative) and Scales (Goal Relevance vs. Requirements) matched the structure of Goals Relevant, Goals Irrelevant, Requirements Must, and Requirements Won't while controlling for answering biases.

I averaged the level of agreement to the items on each scale. I used these measures of agreement to run a 2 (Stakeholders' View: Business vs. Personal) * 2 (Goal Orientation: Egotistic vs. Altruistic) (between-subjects) * 2 (Item Type: Indicative vs. Contra-indicative) * 2 (Scales: Goal Relevance vs. Requirements) (within-subjects) MANOVA. The results are displayed in Figure 1.

I found that the third-order interaction of Item Type (Indicative vs. Contra-indicative) by Scales (Goal Relevance vs. Requirements) by Stakeholders' View (Business vs. Personal) was significant at $\alpha = .05$, $F_{(1,30)} = 10.19$, $p = .003$, $\eta_p^2 = .25$ (for the complete analysis, see Hoorn & Kok, 2005, Tech. Rep. [CD]). The grand mean levels of agreement that underlie this result are depicted in Figure 1. For relevant goals, left of the vertical line, grand mean agreement according to the Personal View (M= 3.92, SD= .76) did not differ much from the Business point of View (M= 3.86, SD= .73). Regarding Irrelevant Goals, the difference between Personal (M= 2.43, SD= .80) and Business (M= .73, SD= .54) was much larger. On the whole, the agreement with irrelevant goals dropped, as could be expected, but the disagreement to irrelevant *business* goals was more severe. In general, the level of agreement with personal goals was higher than with business goals.

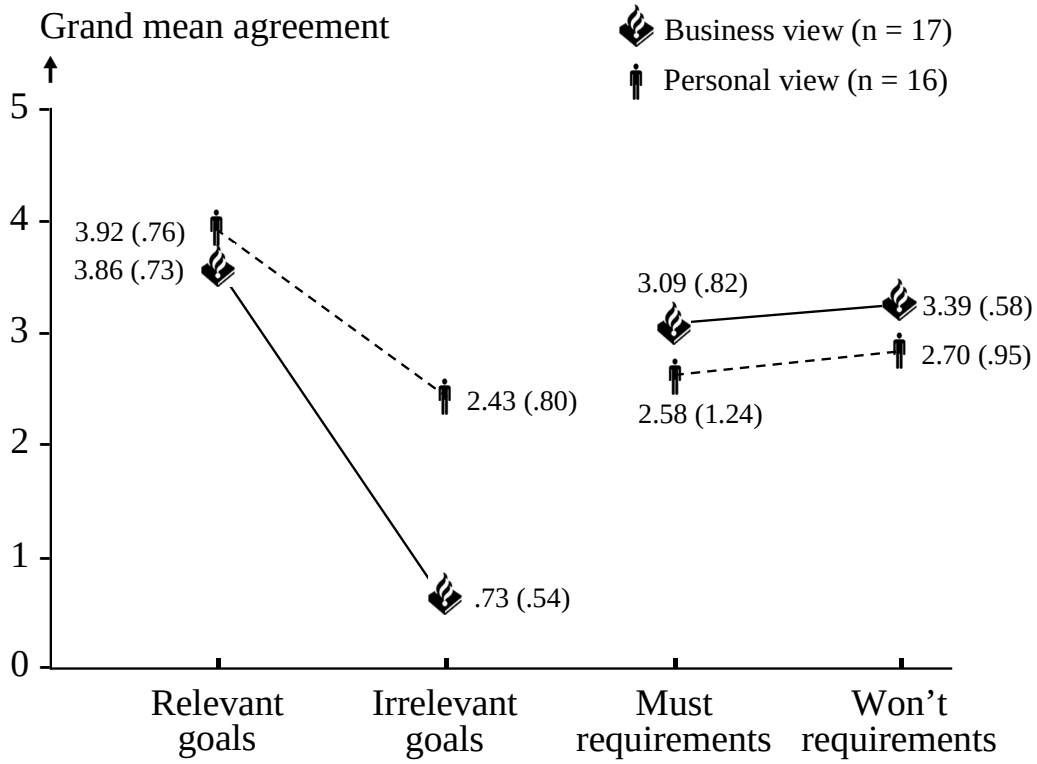


Figure 1: Grand mean averages of agreement to Goal Relevance (Relevant vs. Irrelevant) and Requirements (Must vs. Won't) from a Business and a Personal Viewpoint. Standard deviations are between parentheses (N= 33). The vertical line designates the requirements-analysis rift (see text).

With regard to requirements, however, the pattern changed (right from the vertical line). For both Must and Won't Requirements, the level of agreement from a Business Viewpoint (M_{Must}= 3.09, SD= .82; $M_{Won't}$= 3.39, SD= .58) was always higher than from a Personal point of View (M_{Must}= 2.58, SD= 1.24; $M_{Won't}$= 2.70, SD= .95).

This interaction effect explained a quarter of the variance in agreement (η_p^2= .25), which is considerable. Note that, interestingly, Won't Requirements raised more agreement than Must Requirements and that the factor Goal Orientation (whether goals had a more egotistic or altruistic quality) only yielded insignificant effects (Hoorn & Kok, 2005, Tech. Rep. [CD]). Thus, in refining stakeholders' goals to develop requirements, analysts do not have to ponder the distinction between selfish and social wishes; for now, the more general division between business and personal suffices.

I repeated the analysis with Sex (2), Function (2), Years in Service (4), Years in Present Function (3), and Number of Years at the Academy (3) as

fixed factors and controlled for Age as covariate. None of the effects of these background variables were significant (Hoorn & Kok, 2005, Tech. Rep. [CD]).

3.2.3 *Discussion of the Focus Switch*

Hypothesis H1 predicted that police officers would agree with personal goals more than business goals. Indeed, the officers agreed with personal goals more than business goals, irrespective of relevance within a viewpoint. I further hypothesized (H2) that because the requirements for the CMS were set up by the management to serve business goals, agreement with requirements would be lower from a personal viewpoint than from a business perspective. Indeed again, from a business perspective, officers agreed with the management-viewpoint requirements more than they did from a personal point of view.

However, the confirmation of both hypotheses H1 and H2 complicates requirements analysis and validation. Study 1 informed us that management-viewpoint requirements mainly covered strategic business goals rather than the personal goals of the officers. Yet, although the business goals raised relatively little agreement, the management-viewpoint requirements did raise relatively high levels of agreement. *Thus, the officers agreed more with management-viewpoint requirements when they were framed in a business perspective, although their agreement with the related business goals was lower than that of their personal goals (even when these goals were irrelevant). Once the personal viewpoint was activated, however, the level of agreement to the management-viewpoint requirements dropped.*

This seems like a clear-cut victory for management. You sell them what you want with a nod to the latest mission statement, and they'll buy it. You can see, however, that the won't requirements raised more agreement than the must requirements (Figure 1). In other words, the work floor was less interested in what management wanted for the system than what management did not want. Moreover, the requirements analyst should know that while stakeholders may claim agreement with a set of requirements, they may do so in the 'business mode.' On the work floor, however, once the system is up and running and personal goals become activated, stakeholders may realize that the system does not satisfy their needs, and complaints and change requests will flood your desk.

I coined this finding the 'requirements-analysis rift.' Stakeholders on the work floor regard requirements a business matter, whereas they consider goals (whether relevant or irrelevant) that the system addresses as more private. If management presents a list of requirements to the work floor, there is a great danger that people will more-or-less say they agree because requirements 'are something of the business, not mine.' Once the system is implemented, however, one may discover too late that many change requests occur because it does not work well in practice. During software development, the viewpoints approach can uncover the requirements-analysis rift and initiate timely change requirements.

Yet, one might counter that the officers had a few years in service, which threatens the study's validity. However, as a case in point that managers think up a system without consulting the future users, these officers represented the right population. Sure, users may get used to the peculiarities of a software product but that does not mean they are satisfied or do not want to change its features.

3.3 Study 2: Changes in Requirements Prioritization

Study 1 suggests that change requirements or change requests that result from shifts within and between business and personal models can be detected by taking a viewpoints approach to requirements (cf. Sommerville & Sawyer, 1997). However, we did not yet observe requirements change 'in action.' Do requirements indeed change when goals change within different viewpoints? Do requirements themselves become more or less relevant? In other words, whereas Study 1 focused on the relevance of goals, Study 2 looked into the relevance of requirements. Study 2 also differed from Study 1 in that, in this case, the requirements list was explicitly motivated by a set of related goals.

Study 3 was set up as a replication of Study 2 to retest the special application of the Spearman rho statistic (Section 3.3.1.3) and to solve some methodological issues (Hoorn & Breuker, 2005). In both studies, my master student Mark Breuker and I manipulated the viewpoints on a set of requirements for an e-learning environment to study how changes in business or personal models affect priority change. We chose requirements prioritization because it is useful for selecting the features that must or won't be implemented. In addition, a change in prioritization may be regarded as a form of requirements change. Must requirements become could-be or perhaps even won't-be (or vice versa).

In the literature, requirements are prioritized on several criteria. Moisiadis (2002) provides an instructive overview. In almost all cases, prioritization is based on some sort of relevance (Frijda, 1986) to stakeholder concerns, such as urgency (development time until release) (Lam et al., 1999), effort (the work needed to develop and implement a feature) (ibid.), or cost-value trade-offs (Karlsson & Ryan, 1997).

The criteria used for requirements prioritization involve different types of metrics. Customers score requirements on a rating scale for importance (e.g., 1-least important to 10-crucial) (Lam et al., 1999) and prioritize according to urgency (development time until release) (ibid.) and effort (the work needed to develop and implement a feature) (ibid.). Important methods in requirements prioritization, such as the Analytic Hierarchy Process (AHP) applied to cost-value trade-offs (Karlsson & Ryan, 1997), analyze pair-wise comparisons of a set of requirements. Priority of a feature is based on estimated cost to implement a feature against estimated importance to the stakeholder(s) (ibid). In scoring requirements on ordinal scales, Quality Function Deployment (QFD) (Akao, 1990) tries to align customer requirements with design parameters.

Wiegers Prioritization Model (WPM) is rooted in a combination of AHP and QFD.[3] Wiegers conceives of the relative priorities of features as a combination of benefit to the customer versus the risk for the designer to receive 'punishment' if the feature is left out as well as the relative cost versus the technical risk of implementing the feature (Wiegers, 1999). The Requirements Prioritization Tool (RPT) (Moisiadis, 2002) uses questionnaires and rating scales to assess the stakeholders' attitudes to both business objectives and requirements. For a critical discussion of AHP, QFD, and WPM as well as an explanation of RPT, see Moisiadis (2002).

The methods of prioritizing requirements mentioned above all focus on establishing the requirements of one specific system in one specific organization. These approaches are important to calculate trade-offs between requirements (e.g., Andrews, Runeson, & France, 2004) and to engineer the precise features of a system in a real business case. However, when business models change, requirements analysis must start all over again (Moisiadis, 2002), which may be a tedious and frustrating job. In addition, Gervasi & Zowghi (2005) explain that conventional prioritization methods are quite rigid, often logically oriented, and most importantly, more of a tool for the specialized researcher than a natural way for the stakeholder to prioritize requirements.

This calls for an approach that is both more natural and better at identifying areas of priority changes. Knowledge of what particular types of stakeholder views or business goals are vulnerable to change can bring more focus to requirements reanalysis, i.e. priority change.

Large non-profit and governmental organizations tend to model themselves after commercial businesses. In our university, this is definitely the case. Distance-learning, for example, is gaining ground as a cost-saving measure. We wanted to give the egotistic-altruistic goals factor another try. We envisioned business-model change in a non-profit institute like our university as shifts between goals that focus on the continuity and prosperity of the organization itself (egotistic/selfish business goals – B_e) and goals oriented towards others than the organization (altruistic business goals – B_a). The university has altruistic goals to dispense knowledge to the world, educate people, and improve the quality of life. We predicted similar effects of personal goals (i.e. of those involved in distant learning): whether individuals use an educational system to increase their own personal market value (egotistic/selfish personal goal – P_e) or help other students with their work (altruistic personal goal – P_a) may seriously affect the priorities on the requirements list.

I hypothesized (H3) that changes in business goals predicate changes in the prioritization of requirements. The same may be valid for changes in personal goals. Under different conditions of goal change, stakeholders may attribute different relevance (important - unimportant) to requirements.

[3] http://www.processimpact.com/process_assets/requirements_prioritization_worksheet.xls

I also hypothesized (H4) that prioritization differs when business goals change from a more egotistic (e.g., fire more staff) to a more altruistic nature (e.g., educate students) or vice versa. Similarly, prioritization changes when personal goals change from egotistic (e.g., increase my market value) to altruistic (e.g., help my colleagues) and back again.

In the experiment we conducted, the factor Stakeholders' View indicated whether the concerns of the stakeholders were Business or Personal. The factor of Direction of Goal Change indicated whether the business or personal goals changed from Egotistic to Altruistic or the other way round. Next, we report whether Stakeholders' View and the Direction of Goal Change within business and personal viewpoints altered the priorities of requirements for an e-learning system.

3.3.1 *Method*

3.3.1.1 Participants and Experimental Design

From a database of 1005 science students at the Free University of Amsterdam, 968 students were randomly selected and divided into 4 groups of 242 so that each group contained an equal number of students from the same year and field of study (Breuker & Hoorn, 2004, Tech. Rep. [CD]). Students served as volunteers. For each student, the year of enrollment and field of study was administered. To balance the number of students over studies, only recent years were used. Therefore, students from 1999 and earlier were dismissed. The remaining students were assigned to four conditions of goal change:

Business goals, from selfish to altruistic ($B_{e\rightarrow a}$)

Business goals, from altruistic to selfish ($B_{a\rightarrow e}$)

Personal goals, from selfish to altruistic ($P_{e\rightarrow a}$)

Personal goals, from altruistic to selfish ($P_{a\rightarrow e}$)

Thus, the experiment consisted of a 2 (Stakeholders' View: Business vs. Personal) by 2 (Direction of Goal Change: Egotistic-to-altruistic vs. Altruistic-to-egotistic) between-subjects factorial design.

We developed an online survey for each group to measure requirements priority change under one of the conditions of viewpoint or goal change. The four surveys each contained two pages (Appendix 3.2). Each page presented a text that contained one type of goal (either B_e, B_a, P_e, or P_a), called the Motivation. A 6-point Likert item measured the student's global personal attitude towards the Motivation.

The Motivations were made up by the experimenters without doing proper work floor ethnography. The B_e condition Motivation focused on the university's ambition to play a pre-eminent role in developing the knowledge economy. Therefore, the university wanted to deliver top-qualified students to the

market. To acquire European subsidies, the university aimed to improve student performance by introducing the new e-learning environment.

The B_a condition Motivation emphasized the university's societal responsibility to develop the knowledge economy. Therefore, the university wanted to deliver top-qualified students to the market. To ensure that European subsidies would be beneficial to the students, the university aimed to create a stimulating environment through the new e-learning system.

The Motivation in the P_e condition stated that earlier (but fictitious) interviews among the student population showed that students wanted to profit from the advancing knowledge economy. Therefore, the university wanted to deliver top-qualified students to the market. For the students to get a top job, so the interview results showed, students thought that governmental subsidies should be invested in new technology and learning materials to support the students in their studies. In the interviews, the said new e-learning system was designated as the best alternative.

In the P_a condition, the Motivation declared that – in view of earlier interviews – the students felt responsible for advancing the knowledge economy. The students thought that the money invested in their education should also show some returns to society. Therefore, new technology and learning materials were needed for the students to support one another. In the interviews, the said new e-learning system was designated as the best alternative.

The Motivation and Likert attitude item were followed by a requirements list of 16 features (Appendix 3.2) of a new e-learning system, Didactor (Didactor, 2003), that was to replace the current Blackboard system (Blackboard, 1997-2005). The requirements list was derived from prior analysis of Didactor. Examples of proposed features were: Chat box for students and teachers, Discussion forum, Web mail, Portfolio of personal development, Work team support, etc.

The goals described in the Motivation could be either egotistic/selfish (B_e or P_e) or altruistic (B_a or P_a). If the business goals in the Motivation on the first page were egotistic, then the goals on the second page were altruistically oriented ($B_{e\rightarrow a}$). If the business goals in the Motivation on the first page were altruistically oriented, then the goals on the second page were more egotistic ($B_{a\rightarrow e}$). The same procedure was followed for personal goals ($P_{e\rightarrow a}$ vs. $P_{a\rightarrow e}$). For groups 1 and 2, goals were presented as the university's (fictitious) business goals, which motivated the requirements list. For groups 3 and 4, the goals were personally oriented and introduced to the students as the outcome of a series of (fictitious) interviews among a sample of science students. The requirements remained the same under each condition of goal change. However, the presentation order of requirements from top to bottom was randomized and hence, different for each student and within students, different between the two goal conditions.

3.3.1.2 Procedure

Students received an e-mail message (Breuker & Hoorn, 2004, Tech. Rep. [CD]) containing (1) the announcement that a new Learning Management System, Didactor might replace the current Blackboard system, (2) the notice that a survey would evaluate student needs before the system's introduction, and (3) a request to fill out the survey via a personalized hyperlink. The survey automatically randomized the requirements on the list presented to the students. Students were asked to rank order the requirements for relevance ('importance' on the survey). They were encouraged to work quickly so to avoid ties between the first and second list. Upon survey completion, a short debriefing message appeared. In total, 968 e-mail invitations were sent out, yielding responses from 154 people, of whom 103 ranked both lists. The number of responses in each group varied only slightly. The survey was developed and deployed using the PollPoint online survey tool (PollPoint, v7.6.1).

3.3.1.3 Measurements

In the first prioritization experiment, stakeholders put priority scores (1= top priority, 16= no priority) to requirements in the list under the different conditions of viewpoints and goal change $B_{e\rightarrow a}$, $B_{a\rightarrow e}$, $P_{e\rightarrow a}$, and $P_{a\rightarrow e}$. Moreover, they rated their global attitude to the business goals or personal goals on a single-item 5-point rating scale (0= negative, 4= positive). We approached the problem of calculating changes in the priorities of requirements in four ways (Hoorn & Breuker, 2005).

Measure ($\rho_s 1$)

Priority change (the change in priority of requirements under different goal conditions) was established using Spearman's rho (r_s or ρ_s) (Lowry, 2005). ρ_s is a rank-order correlation-coefficient that analyzes whether a bivariate set of paired rankings correlates by rank sum. ρ_s was calculated on the ranks of priority scores that the participants attributed to the features on the requirements list in $B_{e\rightarrow a}$, $B_{a\rightarrow e}$, $P_{e\rightarrow a}$, and $P_{a\rightarrow e}$. ρ_s was used as the operationalization of priority change, which supposedly reflects an aspect of requirements change. The closer ρ_s approached +1, the higher the agreement between the two sets of ranked features (no priority change). The closer ρ_s approached -1, the higher the disagreement between the two sets of ranked requirements (priority change). We predicted that under the influence of changes in the business goals or personal goals, the change in priorities of features in the requirements list will be large, which is reflected in a high disagreement ($-1 \leq \rho_s < 0$) between the two sets of ranked features in, for example, $B_{e\rightarrow a}$. To calculate this measure, we used the data of 103 students who filled out both priority lists and neglected the data of those who prioritized the first list alone. Eleven students who ranked the features on both lists exactly the same where filtered from the data set (see Measure 2 for rationale).

Measure ($\rho_s 2$)

We suspected that the manipulation might have a weak impact, allowing for more-or-less the same priority scores between lists. In that case, Measure 1 is too insensitive to indicate priority change. For each student, therefore, we filtered out those features that received the same priority score and additionally, used only those features that contributed maximally to ρ_s approaching -1. This was accomplished by calculating the squared differences between the features of both lists and selecting the 10 features with scores closest to $\rho_s = -1$. A minimum of 10 rank order pairs is required for ρ_s's critical values to be allowed between -1 and 1 (Lowry, 2005). This operation seems to be begging the question but this is not so. The features that were filtered out are indeed important to the requirements engineer because they seem to be unaffected by personal or business goal change. If these are the same features throughout a stakeholder group, they reflect stable requirements and need to be implemented in the new system if they are high on the priority list. However, our research question was how changes in personal and business models (i.e. goals) affected priority *change*. Thus, we needed to establish change in order to study the effects of personal and business goals. Probably, on each requirements list there is a subset of features that is not sensitive to goal change and one that is. The latter subset can comprise of different features for different people. We even went so far, in order to establish priority change, to filter 11 students with stable requirements (list 1 and 2 were ranked the same) from the sample.

Measure ($\rho_s 3$)

In an e-mail reply, students complained that they did not see the purpose of ranking the requirements twice – albeit from different viewpoints, suggesting that data obtained with the second requirements lists might have suffered from fatigue and training effects.

Therefore, we calculated the requirement to requirement rank-order total-scores. To do so, we used the data of the 154 students who rank-ordered at least the first list of requirements. Possible data obtained by the second list were discarded. We then computed a rank-order total-score in condition B_e (Business egotistic goals), B_a (Business altruistic goals), P_e (Personal egotistic goals), and P_a (Personal altruistic goals). For each requirement, the sum of rank-order scores was computed across all students in a condition. For example, if B_e had 5 students, who scored 1, 3, 1, 6, 9 to a requirement, the sum total for this requirement was 20. If these 5 students all had given this requirement a score of 1, the sum total was 5. If for all 5 students this requirement was put in the 16th place, the sum total was 80. On the basis of the rank-order total score per requirement (which were between 91 and 576), we then rank-ordered the 16 requirements from the lowest to the highest rank-order total score. Subsequently, we replaced the actual rank-order total score by the

rank order number of their relative position in this general priority list. The requirement with the lowest rank-order total score received a 1 and the requirement with the highest rank-order total score received a 16. The requirement to requirement rank-order total-scores were established by calculating, for each student in a condition, ρ_s between

- B_e (as based on the raw data) and the revised B_a (as based on the rank-order total scores)
- B_a (as based on the raw data) and the revised B_e (as based on the rank-order total scores)
- P_e (as based on the raw data) and the revised P_a (as based on the rank-order total scores)
- P_a (as based on the raw data) and the revised P_e (as based on the rank-order total scores)

In so doing, we compared the requirements priorities of each student under one condition with the general requirements priorities of the whole group of students in another condition. If the students in one group showed the same rank-ordering behavior as the group in general in another condition, no change in requirements priorities under conditions of different personal or business goals occurred. If, on the other hand, the students in one group disagreed with the general rank-ordering behavior of a group in another condition, then change in requirements priorities *did* occur.

Measure (ρ_s4)

We combined the approaches of Measure 2 and 3 to arrive at Measure 4. We followed the same procedure as explained for Measure 3 but we excluded those features that received the same rank-order score on both lists. Here also, we selected the 10 features that guaranteed the maximum approach of ρ_s to -1.

3.3.2 *Analysis and Results*

We ran four separate between-subjects ANOVAs of 2 (Stakeholders' View: Business vs. Personal) * 2 (Direction of Goal Change: Egotistic-to-altruistic vs. Altruistic-to-egotistic) on one of the four measures of change in requirements priorities as based on Spearman's rho (ρ_s1, ρ_s2, ρ_s3, ρ_s4). Year of enrollment, Field of study, whether or not the subject rank ordered both Lists in the survey, and Attitude toward the relevant personal or business goals served as covariates. The only significant result at α= .05 was found for Stakeholders' View with Measure 3 (ρ_s3). The means for ρ_s3 in the four conditions of the experiment are shown in Table 3.

Table 3 illustrates that students who ranked the requirements from the Personal goals viewpoint showed more priority change (Mρ_s= .48, SD= .47) than

students who prioritized the requirements from a Business perspective ($M\rho_s$= .60, SD= .21). This difference was underscored by a significant main effect of Stakeholders' View (Business vs. Personal), $F_{(1, 146)}$= 4.09, p< .05, η_p^2= .03.

It should be noted that the effect is small to medium size (Green & Salkind, 2003, p. 171). To gain more degrees of freedom, we ran another 2*2 ANOVA on ρ_s but this time while excluding the covariates. Results improved, but only slightly: Stakeholders' View (Business vs. Personal), $F_{(1, 150)}$= 4.36, p< .04, η_p^2= .03.

Thus, when requirements were prioritized from a Personal point of View, $M\rho_s$= .48 deviated more from 1 (no change) than from a Business Viewpoint ($M\rho_s$= .60). That does not mean, however, that in the Business View no prioritization changes happened. To check this assumption, a one-sample t-test was performed on ρ_s in the Business View (n= 75) with 1 (no change) as the test value. Indeed, ρ_s in the Business View deviated significantly from 1, $t_{(74)}$= -16.36, p= .000. Taken together, the results suggest that some changes in the prioritization of requirements occurred in the Business View but that in the Personal View, this difference was significantly larger.

Table 3: Mean ρ_s for Stakeholders' View and Direction of Goal Change (N= 154).

Stakeholders' View	**Direction of Goal Change**	**Mean ρ_s**	**SD**	**n**
Business	Egotistic to altruistic	.59	.18	36
	Altruistic to egotistic	.61	.24	39
Personal	Egotistic to altruistic	.56	.39	43
	Altruistic to egotistic	.40	.55	36

3.3.3 *Discussion of the Prioritization Experiment*

Study 2 showed that more prioritization activities occurred when personal goals were changing than when business goals were. This sustained H2 (less agreement to requirements from a personal view) in combination with H3 (goal change predicates priority change). Similar to Study 1, stakeholders found their personal goals more important than the business goals, perceived that the requirements did not match their personal concerns, and changed the rank order of requirements on the basis of the requirements' relevance to their personal goals. Also in line with Study 1, the division between egotistic/selfish and altruistic goals was unsubstantial (rejecting H4). Thus, the important thing in Study 2 was *that* goals changed, whether egotistic or altruistic. In other words, if a philanthropic foundation evolves into a commercial enterprise,

changes in requirements prioritization will probably not be due to the switch from altruistic to selfish goals. The mere fact that any kind of goals change (from old to new) will induce prioritization change.

This effect was established using Measure 3 ($\rho_s 3$). Measure 3 was the best among the four measures developed because it was the only measure that did not use the data of the second requirement lists and thus, was not affected by fatigue and training. Another strongpoint of Measure 3 was that it used the data of more students (N= 154), not only those who ranked both requirements lists (n= 103). This by no means released us from the burden of verifying the measure's validity in a replication study, which we did in Study 3.

Further, the effect size of the prioritization change was not very large (η_p^2 = .03), perhaps due to our manipulation. To create the different viewpoints, we developed the goals ourselves rather than gather them on the work floor. Unlike Study 1, then, the relevance of goals in Study 2 was undetermined. Therefore, we devised another requirements-prioritization experiment in a real business setting where the goals would be more appropriate and relevant to stakeholder concerns.

3.4 Study 3: Changes in Requirements Prioritization Revisited

While Study 1 focused on the relevance of goals and Study 2 on the relevance of requirements, Study 3 combined the two and investigated priority change as a function of changes in relevant goals. This also gave us the opportunity to retry our application of the ρ_s statistic and attempt replicating the results obtained in Study 2.

I approached an IBM buy-out, the financial consultants company ConQuaestor, while it was in the process of transforming its consultants training system.[4] The business-model change was reflected in the change of the training system. Previously, ConQuaestor served clients based on the availability of their consultants. However, when consultants were training in the classroom, they were unavailable to their clients. Due to a changing market, ConQuaestor found that their clients, who required financial expertise on demand at any time, were gaining power over consultant availability. As a result, consultants were frequently 'forced' by their clients to skip the training. This negatively impacted the overall knowledge value of ConQuaestor and constrained consultants' ability to take courses (gain new knowledge), which was perceived as a part of their salary. To solve this problem, ConQuaestor currently is in the process of introducing online e-learning courses, accessible to the consultants at any time, anywhere. This should support on-the-job training to relieve the consultants from travelling to and from training locations and allow consultants to better accommodate their clients' needs.

That ConQuaestor moved from an old business model of business-oriented, inflexible training to a new, flexible, customer-oriented model also

[4] http://www.conquaestor.nl/

offered Mark Breuker and me the opportunity to replace the failing hypothesis H4 on egotistic and altruistic goals with something more appropriate to the situation. H5 predicted that business models as well as personal models changing from old to new goals or vice versa (factor Model Change) would seriously impact the prioritization of requirements for, in this case, the new e-learning environment.

3.4.1 *Method*

3.4.1.1 Participants and Experimental Design

I started the investigations by interviewing the top management about the new business goals they wanted to achieve with the e-learning system (e.g., flexibility, on-demand, customer oriented). I also wanted to know the old business goals they wanted to leave (e.g., business oriented, fixed schedules). My research assistant, Mark Breuker, further interviewed a number of consultants to acquire the new personal goals they wanted to achieve with the system (e.g., cognitive growth) and asked what goals they left behind (e.g., disrupting the work to do the training).

In a pretest, the four different sets of goals were scored for relevance by means of 6-point Likert items. Based on Corrected Item-Total Correlations and standardized Cronbach's alpha, we selected six goals that best measured relevance within a condition (Hoorn & Breuker, 2005, Tech. Rep. [CD]). We then created four Motivations based on these most relevant goals per condition (see below).

From a database of 268 ConQuaestor consultants, 206 that were not interviewed during the preparations for the main experiment were selected. These consultants were randomly distributed over 4 groups so that two groups contained 52 and two groups contained 51 consultants. The consultants served as volunteers. Due to privacy reasons, we were not allowed to administer demographics questions, except for e-mail addresses. The groups were assigned to 4 different conditions of viewpoints and goals:

1. Business View, Old Goals (B_o)
2. Business View, New Goals (B_n)
3. Personal View, Old Goals (P_o)
4. Personal View, New Goals (P_n)

Thus, the experiment consisted of a 2 (Stakeholders' View: Business vs. Personal) by 2 (Goal Type: Old vs. New) between-subjects factorial design. For each group, a small online survey was developed to gather the prioritization data. The four surveys each contained a single page (Appendix 3.3). Each page presented a set of six learning-related goals (either B_o, B_n, P_o, or P_n), called the Motivation. The consultants were asked to read the Motivation and indicate how relevant each of these goals were to them on a 6-point rating scale (1= very unimportant, 6= very important). In all four conditions the Mo-

tivation started with a short description of ConQuaestor's general ambitions to be a financial knowledge organization strongly focused on developing skills and knowledge of its employees.

In the B_o condition, the Motivation described the traditional classroom training-situation in which consultants were scheduled to follow training courses at an external location. We emphasized that ConQuaestor strived for an efficient plan that would satisfy the knowledge requirements of the consultants as well as client demands. Due to the classroom setting, however, the training sometimes did not entirely match the individual consultant's level of knowledge. Following this description was a statement that ConQuaestor valued the development of their consultants' knowledge. Therefore, the effectiveness of the training system was under review, focusing on how the training system would support the learning goals specified by the ConQuaestor management. Next the six old business goals were introduced and the consultants were asked to rate them for relevance. They prioritized the requirements thereafter.

In the B_n condition, the Motivation emphasized the need for an e-learning system that would facilitate knowledge on demand for ConQuaestor consultants. This would help the consultants to gain fast and easy access to new knowledge and improve service for the client. The new system would help ConQuaestor to maintain their current training budget while providing training that matched the individual consultant's knowledge level. Following this description was the announcement that ConQuaestor valued the development of their consultants' knowledge and that the new system should support the training-related goals specified by the ConQuaestor management. Next, the six new business goals were presented, and the consultants were asked to rate them for relevance, after which they prioritized the requirements.

The Motivation in the P_o condition was similar to the B_o condition except that the focus was more on the personal aspects of classroom training, such as personal contact with colleagues and with the instructor. The Motivation again stated that it was hard to tailor the supplied course contents to the individual consultant's needs. Again, ConQuaestor attached great value to the consultants' knowledge level, and the effectiveness of the classroom training-system was under review. The aim was to see how the personal learning goals of the consultants were supported by the training system. Next, six old personal learning goals as based on the interviews with the consultants' colleagues were presented and the consultants were asked to rate them for relevance. Prioritization of requirements followed.

In the P_n condition, the Motivation stated – as in the B_n condition – the need for an e-learning system to facilitate learning on-demand. The emphasis was on personal aspects of the new system such as a more flexible support for the learning demands of ConQuaestor consultants and support for one's personal learning style. Because new knowledge would be accessible at any time and place, the consultants would be able to help their clients in a better way. Similar to the other conditions, this description was followed by the statement that Con-

Quaestor valued the development of their consultants' knowledge and that the new system should support their personal learning goals. Next, the six new personal learning goals (based on the interviews with colleagues) were presented and the consultants were asked to rate them for relevance, after which they prioritized the requirements.

The Motivations and goal items were followed by the same requirements list of 16 requirements. The requirements were written in such a way that they could belong to a classroom training-system as well as an e-learning system. Examples of the proposed features were "Short duration courses about a single subject," "Knowledge documentation and sharing," "Trainees follow a predetermined course schedule," etc. (Appendix 3.3). To make sure the consultants would be able to distinguish and rank-order the requirements, each requirement had a counterpart, that is, a competitive requirement. For instance, "Trainees may collaborate to help each other" was accompanied by "Trainees should work on their own."

Because we learned from Study 2 (also Hoorn & Breuker, 2005) that respondents did not like to rank order the same set of requirements twice (albeit under different conditions), each consultant rank ordered the list of requirements only once. Prioritization change could then be measured by comparing the rank ordering results of an individual consultant in one condition with the rank-order total-score of the group in a related condition (Sections 3.3.1.3 and 3.4.1.4).

3.4.1.2 System

ConQuaestor tried one online learning solution, experimentally offered to a small group of thirty consultants. However, due to technical imperfections of the system, it was scarcely used, and consultants rated it poorly. Consequently, this system was discarded, and ConQuaestor is currently looking for a better performing e-learning solution.

3.4.1.3 Procedure

The procedure was the same as described in Section 3.3.1.2 (also Hoorn & Breuker, 2005, Tech. Rep. [CD]). In total, 206 e-mail invitations were sent out, yielding a response from 106 people (51%). The number of responses per group follows: B_o, n = 25; B_n, n = 28; P_o, n = 33, and P_n, n = 20. The survey was developed and deployed using the PollPoint online survey tool (PollPoint, v7.6.1).

3.4.1.4 Measurements

Participants assigned priority scores (1= top priority, 16= no priority) to a list of 16 requirements under the conditions B_o, B_n, P_o, and P_n. Moreover, they rated 6 goals related to their condition for relevance on a 6-point rating scale (1= very unimportant, 6= very important).

The calculation of priority change by means of ρ_s was similar to that in Section 3.3.1.3. Thus, the rank ordering of requirements by individual consult-

ants in condition B_o was compared with the rank-order total-score of requirements of the whole group in the B_n condition. Scores of individuals in B_n were compared with the rank-order total-score in B_o. Likewise, individuals in P_o were compared with the group scores in P_n and vice versa. This way, we created the Model Change (Old-to-new vs. New-to-old) factor used in the analysis of effects on ρ_s.

3.4.2 *Analysis and Results*

3.4.2.1 Prioritization Analysis

From the original 106 cases, we removed 9 extremes and outliers (casewise) of the ρ_s statistic, which left us with N= 97 (55 male, 32 female, 10 unidentified).[5] We then ran a 2 (Stakeholders' View: Business vs. Personal) by 2 (Model Change: Old-to-new vs. New-to-old) (between-subjects) ANOVA on the mean ρ_s, following the calculation in Section 3.3.1.3. However, none of the effects were significant at α= .05.

Yet, careful inspection of the data unveiled that within conditions (i.e. Business Old-to-new) certain consultants were more actively changing priorities than others were. I therefore devised an extra factor of Change Sensitivity (Change Prone vs. Less Change Prone) by performing a median split on the mean ρ_s value (median= .57, cumulative percent= 50.6%). The idea was that consultants with ρ_s greater than the median were inherently less prone to change priorities than consultants with ρ_s smaller than the median, who supposedly were more eager to change. Subsequently, I ran another (between-subjects) ANOVA of 2 (Stakeholders' View: Business vs. Personal) by 2 (Model Change: Old-to-new vs. New-to-old) by 2 (Change Sensitivity: Change Prone vs. Less Change Prone) on the mean ρ_s value (Figure 2). Reanalyses with Sex as fixed factor and Relevance of Goals as covariate did not change the results.

[5] Identification of the sexes was based on the first names used in the e-mail addresses.

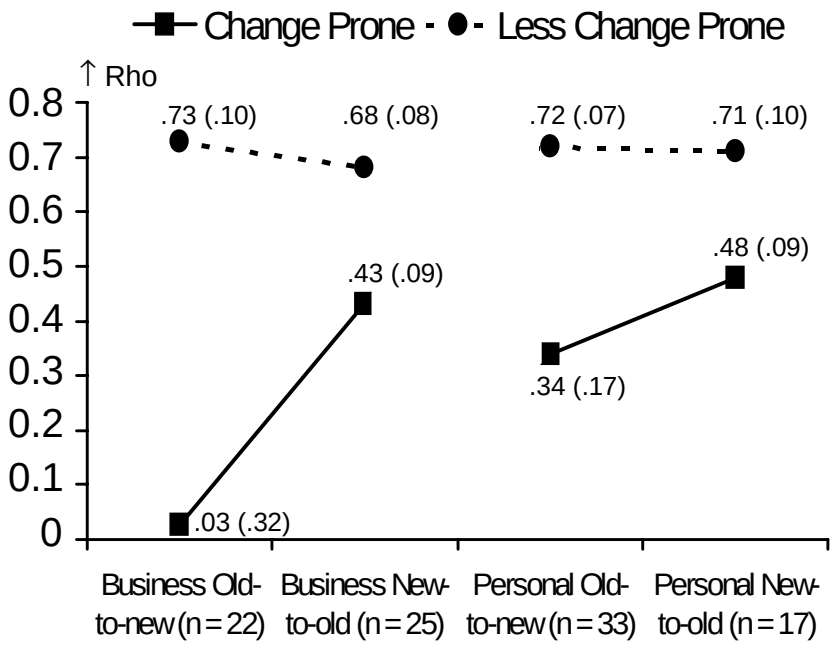


Figure 2: Mean ρ_s value for changes from Old-to-new (and v.v.) Business or Personal models for consultants with more or less sensitivity to changing situations. Standard deviations are in parentheses (N= 97).

The main effect of Stakeholders' View, $F_{(1,89)}$= 11.10, p= .001, η_p^2= .11, repeated Study 2's finding (supporting H2 and H3) that in general, stakeholders changed priority scores more when they prioritized requirements from a Personal View (Mρ_s= .53, SD= .22) than from a Business Viewpoint (Mρ_s= .56, SD= .26). Although in an absolute sense, the difference between the values was small, it was a stable effect with a considerable effect size (η_p^2= .11) (Green & Salkind, 2003, p. 171), much larger than found in Study 2 (η_p^2= .03).

However, the effect of Stakeholders' View was modulated by a number of other factors. The 3-way interaction of Stakeholders' View, Model Change, and Change Sensitivity, $F_{(1,89)}$= 6.02, p= .016, η_p^2= .06, showed that stakeholders could be divided into two groups (Figure 2). The first group consisted of people who were stable in their prioritization under any change of personal or business model. The second group included those who were more sensitive to change and who were inclined to prioritize requirements differently when personal or business models changed. This divide was also supported by a strong main effect of Change Sensitivity, $F_{(1,89)}$= 183.57, p= .000, η_p^2= .67.

Particularly in the group that was change prone, the prioritization differences were larger when the change of models went from Old-to-new than v.v. The strong interaction between Model Change and Change Sensitivity, $F_{(1,89)}$=

26.93, p= .000, η_p^2= .23 also demonstrates this. In other words, these consultants wanted things to change more radically when they prioritized the requirements from an 'old' point of view than when they did so from a 'new' point of view (supporting H5). This was especially true for change-prone consultants who worked within the old business model as compared to the new (Figure 2). Here, ρ_s approached almost zero (Mρ_s= .03, SD= .32), meaning that the rank ordering of the two lists was almost unrelated, that is, the prioritization differed almost completely between old and new (H5). Change-prone consultants were less inclined to change the priorities of requirements when they worked from a new personal model (Mρ_s= .48, SD= .09) as compared to the old (H3 in combination with H5). Thus, these stakeholders seem most satisfied with their new personal 'state of mind' and least satisfied with the old business situation.

3.4.2.2 Relevance of Goals

Apart from the prioritization scores, I also analyzed the effects of viewpoints and sensitivity to change on the level of relevance of goals. On the data of all 106 stakeholders, I performed a (between-subjects) ANOVA of 2 (Stakeholders' View: Business vs. Personal) by 2 (Goal Type: Old vs. New) by 2 (Change Sensitivity: Change Prone vs. Less Change Prone) on the mean Relevance of goals.

A significant main effect of Stakeholders' View occurred, $F_{(1,98)}$= 26.93, p= .000, η_p^2= .23, sustaining the findings in Study 1 (H1) that on the whole, the relevance of personal goals (M= 4.49, SD= .48) was higher than that of business goals (M= 4.07, SD= .46). However, two significant interactions refined this view.

The interaction between Stakeholders' View and Goal Type, $F_{(1,98)}$= 10.26, p= .002, η_p^2= .10, indicated that the level of relevance of Old Goals in the Personal View (M= 4.35, SD= .50) was lower than to New Goals in the Personal View (M= 4.73, SD= .34). This may be expected in view of the prioritization results. Yet, the level of relevance of Old Goals in the Business View (M= 4.15, SD= .44) was *higher* than of New Goals in the Business View (M= 4.00, SD= .46).

This position can be understood by scrutinizing the significant interaction between Stakeholders' View and Change Sensitivity, $F_{(1,98)}$= 3.97, p= .049, η_p^2= .04. In the Business View, the consultants Less Prone to Change (M= 3.98, SD= .47) attributed less relevance to business goals (whether old or new) than those who were more Change Prone (M= 4.18, SD= .41). However, from a Personal Viewpoint, those Less Prone to Change (M= 4.57, SD= .46) attributed more relevance to their personal goals (whether old or new) than consultants who were more willing to change, Change Prone (M= 4.44, SD= .50). In plain words, the more conservative consultants (less prone to change) saw less relevance in any type of business goal than more liberal consultants (those

more willing to change). Conversely, conservatives esteemed their own personal goals (old and new) higher than the liberals their own liberal goals.

3.4.3 *Discussion of Prioritization, Old and New Goals, and Conservatism*

The more conservative consultants probably account for the effect that old business goals were more relevant than new ones. These people valued their own goals higher than other goals, and no business change could influence that. Therefore, they did not feel the need to change priorities of requirements. Business goals as well as system requirements could stay as they were. That nevertheless the prioritization of requirements changed the most in the old business situation may be the work of the more liberal forces (see Figure 2). Compared to the conservatives, they valued goals other than their own as well and were less convinced about their own goals than conservatives were. Put into the old business situation, they were therefore willing to set new personal goals and seriously change the priorities on the requirements list.

The upshot is that part of the said situation may be attributed to men's versus women's attitudes. As a control factor, we added Sex to our analysis and found that the level of agreement to Goal Type (Old vs. New) was affected by Sex, $F_{(1,78)}= 6.64$, $p= .012$, $\eta_p^2= .08$. Women agreed more with the New Goals ($M= 4.57$, $SD= .45$) and disagreed more with the Old Goals ($M= 4.09$, $SD= .51$) than men, who agreed with the Old Goals more ($M= 4.34$, $SD= .47$) than the New ($M= 4.15$, $SD= .54$). In other words, the men were the more conservative consultants whereas the women probably contributed strongly to the more liberal forces, advocating a change of business as well as software requirements.

In Study 3, the introduction of online education and training into the company was a major issue, and it could be argued that this was more important than the requirements in terms of features of the proposed system. The overshadowing of the system by this issue may have threatened the validity and results of Study 3. Study 2, in contrast, seemed more secure in terms of validity. Here, the idea of online study should not have been an issue for students, as they were already using Blackboard. They should know and appreciate the value of the requirements proposed. Their viewpoint could be easily distinguished from that of the University, and they after all were the ones who had to use the system.

Nonetheless, introducing e-learning to ConQuaestor was the concrete manifestation of a business-model change, which we captured by devising the factor Model Change (Old-to-new vs. New-to-old). As argued, H5 predicted prioritization changes as a function of Model Change, which indeed happened (the prioritization changes were bigger from old to new). It is reassuring that the effects of Stakeholders' View (Business vs. Personal) in Study 2 could be repeated in Study 3. Because Study 2 was less problematic regarding validity, repeating (and improving) its effects in Study 3 seem reason to trust the validity of Study 3 (cf. Ohlsson & Runeson (2002).

3.5 General Conclusions and Discussion

Requirements change (e.g., Highsmith & Cockburn, 2001) and stakeholders seem inconsistent about what they want (Van Lamsweerde, 2004). Still, we can understand this better from a viewpoints (Sommerville & Sawyer, 1997) and goal-oriented position. Relevance in particular plays a major role in the occurrence of change requirements. A different viewpoint causes different goals to become relevant. This affects the relevance of requirements for these goals, which can be made visible in the priority list (cf. Lam et al., 1999). That is, fluctuations in the level of relevance of requirements to changing goals are viewpoint-dependent and lead to different prioritizations.

3.5.1 *It's Okay for the Business but not Necessarily for Me*

The most important finding in this Chapter is the existence of a requirements-analysis rift. Stakeholders consider requirements as the business's concern, but it is most important that these requirements meet personal goals. Yet, stakeholders do not necessarily connect management-viewpoint requirements to their personal goals because that requires changing their focus from a business to a personal point of view (Figure 3).

If in practice, a group of stakeholders agree upon requirements, be sure that they used the right perspective. That is, viewpoints on requirements and goals should be the same *and* congruent with the stakeholders' own point of view. Otherwise, it becomes very hard to interpret how the stakeholders actually agree. There are at least eight possibilities. The first two bullet points form the rift as established in the CMS case.

Workforce judging

- Business requirements linked to business goals
- Business requirements linked to personal goals
- Personal requirements linked to business goals
- Personal requirements linked to personal goals

Likewise for the managers.

This bulleted list shows that, in the future, it would be worthwhile to explore the distinction between managers and officers while they take different viewpoints, expressing agreement to different types of requirements that are consequential to different types of goals. Another relevant viewpoint, not explored in Study 1, would be that of the clerical officers who have to do the scheduling using the CMS.

In this respect, the requirements-analysis rift is a novel finding and a research framework that can account for Davis's (1990) point that in practice the need for traceability from goals to requirements is germane (p. 193). Our re-

sults imply that, first, the practitioner should clarify for future end-users what the consequences of their tasks and work processes are once a system as demanded by the management is operational. Second, in the requirements specification, it should be possible to trace back functionality to business as well as to personal goals.

I tested five hypotheses, and the results support and refine the rendition of the requirements-analysis rift. H1 expected that personal goals are regarded as more important than business goals. This was indeed the case in two studies that investigated the relevance of goals. In Study 1, officers who judged the management-viewpoint requirements on a Capacity Management System agreed more with their personal goals than the business goals, even when these personal goals were irrelevant from the personal viewpoint. In Study 3, consultants who judged the requirements on an e-learning system deemed their personal goals more relevant than the business goals. This was particularly the case for male consultants who were less open to change. Their personal goals prevailed throughout business-model change. Females who were more open to change attached somewhat more relevance to the business goals than the conservative males, but these liberal females regarded their personal goals even more important.

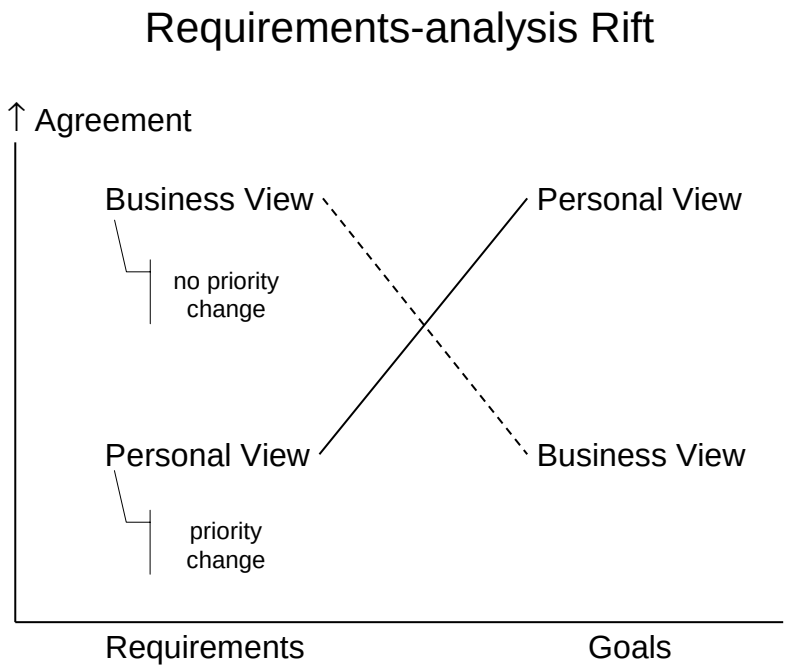


Figure 3: The requirements-analysis rift occurs when goals and requirements are perceived as unconnected lists.

The counterpart of the requirements-analysis rift indicates that requirements are considered a manifestation of the business. This point is addressed by H2,

expecting agreement with requirements to be higher from a business than from a personal viewpoint. All three studies confirmed this idea with different stakeholders (i.e., police officers, university students, and financial consultants). Study 1 showed that officers agreed more to both must and won't requirements from a business viewpoint than from a personal perspective. Study 2 with students and Study 3 with consultants both showed that stakeholders were more eager to change the priorities of requirements when their personal situation changed than when the business changed. Thus, stakeholders more-or-less left the priority list unchanged from a business viewpoint but disagreed with the prioritization and adjusted what was important once they took a personal stand. When stakeholders were satisfied with their current personal situation, they were less inclined to change the prioritization even when by default they were willing to change.

H3 claimed that whether business or personal, goal change predicates priority change. This was indeed the case in Study 2 and 3. In Study 2, students changed the priorities of requirements when the business model changed but rearranged the priorities even more when personal models were altered. The same happened in Study 3 with some modifications. Stakeholders who valued their personal goals highly and were less willing to change did not comply with H3. Their prioritization remained stable through any goal changes. Those more open to change did confirm H3, and changes in priorities were even greater when individuals used business goals while themselves adopting a new set of personal goals.

I also looked into the type of goals that could change. H4 believed that changes from egotistic goals (e.g., become market leader) to altruistic goals (e.g., cure patients) or v.v. impinges upon the prioritization of requirements. This possibility was rejected, however. Study 1 and Study 2 showed that stakeholders did not distinguish between more self-centered or more socially oriented goals, neither for agreement to goals, agreement to requirements (Study 1), nor for the prioritization of requirements (Study 2). Thus, goal changes had an effect (H3) but goals being more selfish or more altruistically oriented did not (refuting H4).

Instead, the changes between old and new personal or business goals did make a difference. H5 affirmed that whether personal or business, changes from an old model to a new or from a new model back to an old impacted the priorities of requirements. Study 3 showed that this was only valid for stakeholders with an a priori sensitivity to change. Conservative stakeholders did not prioritize the requirements differently in old or new situations because they adhered to their own personal goals throughout. They showed the least priority change when they worked from an old business model as compared to a new. Stakeholders who were most sensitive to change showed exactly the opposite pattern. When *they* worked from an old business model and we compared it to what stakeholders did in a new business model, they wanted to change the priority of almost every requirement on the list. As a group, change-prone

stakeholders were less willing to alter the priorities when they worked from a new personal model and we compared it to an old personal model. These findings may suggest a solution to bridge the rift (Figure 4).

Figure 4 is applicable only to stakeholders who are willing to change, but these people are exactly the ones to put a change requirement. First of all, requirements should be connected to goals. The Motivations we wrote in Study 3 were based on the most relevant goals stakeholders could think of in relation to the to-be-developed system. It turned out that change-prone stakeholders – the ones that prompt change requirements easily – are most satisfied and less willing to change when requirements connect to their newly gained personal insights and goals. When they are allowed to compare these new personal goals with the old business model (the thing most distant from their new personal situation) and then judge the requirements from their personal point of view (Figure 4, solid), the urge to change the requirements will drop as will the frequency of change requirements. In contrast, if they work from the old business model while themselves adhering to newly acquired goals (Figure 4, dashed), they will change almost every requirement that's on the list, the same list as used before, that is. It is all a matter of framing.

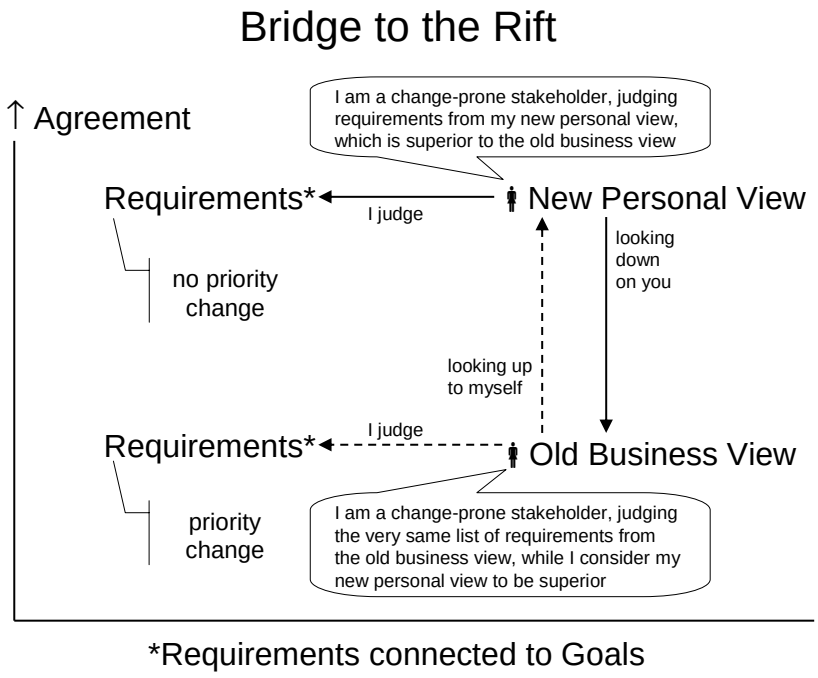


Figure 4: The requirements-analysis rift can be bridged when goals and requirements are connected and stakeholders can compare between old business and new personal situations.

3.5.2 *Message to Practitioners*

You may go to the work floor, ask what they want, go back to design a system, and have them judge the requirements to which they more-or-less agree only

to find out that after the system is delivered, the work floor complains that your system does not serve their purposes. You may have stumbled into the requirements-analysis rift. That is, people can hardly imagine that requirements will impact their future work situation. Requirements are considered part of the business, serving business goals. That's why they agree to them; the requirements you offered serve the business goals well. Only if you can connect management-viewpoint requirements to personal goals (particularly newly acquired goals), will stakeholders tell you what they really think. Thus, do not work from loose requirements lists. Remember that attending to personal goals is more important for making a software development-project a success (cf. Cooper & Reimann, 2003) than looking after business goals (which are also important but to a lesser extent).

Volatile requirements and volatile stakeholders exist. You can detect both with the Spearman rho (ρ_s) application we developed. For example, if you have a list of requirements rank ordered on two different moments during a software-development track, the requirements that contribute most to ρ_s= -1 are the ones most likely to become a change requirement. You can do this by calculating the squared differences between the requirements of both lists and identify which requirement has scores closest to ρ_s = -1 (Hoorn & Breuker, 2005). The requirements that contribute most to ρ_s= 1, are the stable requirements. If they moreover are high on the priority list, chances are high that you can readily implement these.

By doing a median split on the mean ρ_s value, you can identify which stakeholders are more eager to change the requirements than others. Those more ready to propose a change requirement are below the median (they produce values closer to ρ_s= -1). Those stakeholders who are stable throughout changing situations are above the median, producing values closer to ρ_s= 1.

Taken in unison, the danger zone of change requirements in any development project are the requirements that come closest to ρ_s= -1 of stakeholders below the median. That is, the things liberal people want to change most. The safe zone are the requirements that contribute most to ρ_s= 1 of stakeholders above the median. That is, the things conservative people don't want to change at all. Not all people change their minds all the time.

References

Akao, Y. (1990). *Quality Function Deployment QFD: Integrating Customer Requirements into Product Design.* Cambridge, MA: Productivity Press.

Alexander, I., & Robertson, S. (2004). Understanding project sociology by modeling stakeholders. *IEEE Software, 21, 1,* 23-27.

Andrews, A. A., Runeson, P., France, R. B. (2004). Requirements trade-offs during UML design. *Proceedings of the 11th IEEE International Conference and Workshop on the Engineering of Computer-Based Systems (ECBS'04)* (pp. 282-292), May 24-28, Brno, Czech Republic. Los Alamitos, CA: IEEE Computer Society.

BISON (2003). Mass confusion in telecom world paves the way for booming consulting franchise. Altruistic business consultants find recession-proof security in Schooley Mitchell. *Bizjournals* (January 3). Retrieved Sept. 21, 2005 from http://bizjournals.bison. com/press/pr1-3-03schooley.html.

Blackboard (1997-2005). Blackboard Inc., WA.

Boehm, B. W., & Pappacio, P. N. (1988). Understanding and controlling software costs. *IEEE Transactions on Software Engineering, 14,* 1462-1477.

Breuker, M. E., & Hoorn, J. F. (2004). *The E-learning Hoax: Didactor versus Blackboard.* (Tech. Rep.) [CD]. Amsterdam: Free University.

Cockburn, A., & Highsmith, J. (2001). Agile software development: The people factor. *IEEE Computer, 34, 9,* 131-133.

Cooper, A., & Reimann, R. (2003). *About Face 2.0: The Essentials of Interaction Design.* New York, NY: Wiley.

Davis, A. M. (1990). *Software Requirements - Analysis and Specification.* Englewood Cliffs, NJ: Prentice Hall.

Didactor (2003). The Mediator Group, Amsterdam.

Dillman, D. A. (1999). *Mail and Internet Surveys: The Tailored Design Method.* New York, NY: Wiley.

Easterbrook, S., & Nuseibeh, B. (1996). Using ViewPoints for inconsistency management. *BCS/IEEE Software Engineering Journal, 11, 1,* 31-43

eRA (2002). Program Portal Scope Document v1.4. (Tech. Rep.). National Institutes of Health, Bethesda MD. http://era.nih.gov /Docs/Pgm_Por_Scope_4-02.doc

Frijda, N. H. (1986). *The Emotions.* New York, NY: Cambridge University Press.

Gervasi, G., & Zowghi, D. (2005). Reasoning about inconsistencies in natural language requirements. *ACM Transactions on Software Engineering and Methodology, 14,* 277–330.

Green, S., & Salkind, N. (2003). *Using SPSS for Windows and Macintosh: Analyzing and Understanding Data.* Upper Saddle River, NJ: Prentice-Hall.

Highsmith, J., & Cockburn, A. (2001). Agile software development: The business of innovation. *IEEE Computer, 34, 11,* 120-122.

Hoorn, J. F., & Breuker, M. E. (2005). Panta rho: Personal before business in changing requirements priorities. In J. Ralyté, P. J. Ågerfalk, & N. Kraiem (Eds.), *Proceedings of the First International Workshop on Situational Requirements Engineering Processes: Methods, Techniques, and Tools to Support Situation-Specific Requirements Engineering Processes (SREP'05), Paris, France, August 29-30, 2005, In Conjunction with the Thirteenth IEEE Requirements Engineering Conference (RE'05)* (pp. 152-162). Ireland: University of Limerick.

Hoorn, J. F., & Breuker, M. E. (2005). *Requirements Prioritization at a Financial Buy-out of IBM.* (Tech. Rep.) [CD]. Amsterdam: Free University.

Hoorn, J. F., & Kok, E. (2005). *Requirements Validation of a Capacity Management System at the Dutch Police Force.* (Tech. Rep.) [CD]. Amsterdam: Free University.

Hopkins, J. (2002). Minority women go it alone in starting their own businesses. *USA Today* (October 22). Retrieved Sept. 21, 2005 from http://www.usatoday.com/educate/Entre11.pdf.

Karlsson, J., & Ryan, K. (1997). A cost-value approach for prioritizing requirements. *IEEE Software, 14, 5,* 67-74.

Lam, W., Shankararaman, V., & Saward, G. (1999). Requirements change: A dissection of management issues. *Proceedings of REFSQ,* Heidelberg, Germany.

Lehman, M. M. (1996). Laws of software evolution revisited. Lecture notes in Computer Science, 1149. *Proceedings of the 5th European Workshop on Software Process Technology*, London, UK.

Leite, J. C. S. P., & Freeman, P. A. (1991). Requirements validation through viewpoint resolution. *IEEE Transactions on Software Engineering, 17, 12*, 1253-1269.

Loucopoulos, P., & Kavakli, E. (1995). Enterprise modelling and the teleological approach to requirements engineering. *International Journal of Cooperative Information Systems, 4, 1,* 45-79.

Lowry, R. (2005). Spearman rank order correlation coefficient. *VassarStats.* http://faculty.vassar.edu/lowry/corr_rank.html. Accessed August 12, 2005.

Moisiadis, F. (2002). The fundamentals of prioritising requirements. *Proceedings of SETE,* Sydney, Australia.

Ohlsson, M. C., & Runeson, P. (2002). Experience from replicating empirical studies on prediction models. *Proceedings of the 8th IEEE International Symposium on Software Metrics (METRICS'02),* Ottawa, Canada (pp. 217-226). Los Alamitos, CA: IEEE Computer Society.

PollPoint (V7.6.1). *PollPoint Online Survey's.* Retrieved July 1, 2004 from http://www.pollpoint.nl.

Potts, C., & Takahashi, K. (1993). An active hypertext model for system requirements. *Proceedings of the 7th International Workshop on Software Specification and Design*, Redondo Beach, CA.

Price, J., & Cybulski, J. (2004). Influence of stakeholder communication on consensus making in requirements negotiation. In *Proceedings of AWRE,* Adelaide, Australia.

Sommerville, I., & Sawyer, P. (1997). Viewpoints: Principles, problems and a practical approach to requirements engineering. *Annals of Software Engineering, 3*, 101-130.

Strens, M. R., & Sugden, R. C. (1996). Change analysis: A step towards meeting the challenge of changing requirements. *IEEE Symposium and Workshop on Engineering of Computer Based Systems (ECBS'96)*, Friedrichshafen, Germany.

Van Lamsweerde, A. (2004). Goal-oriented requirements engineering: A roundtrip from research to practice. *Proceedings of RE'04,* Kyoto, Japan.

Wiegers, K. E. (1999). *Software Requirements.* Redmond, WA: Microsoft Press.

Appendix 3.1

CMS RE*quest*, version P_e

Vragenlijst Capaciteit Management Politie: Groep P_e

De volgende stellingen gaan over de doelen die u persoonlijk in uw werk heeft.
U geeft uw mening door één van de getallen 0 t/m 5 te omcirkelen.

Persoonlijke doelen bij het werk
Ik vind het belangrijk dat…

PE06i ik gewaardeerd word door mijn chef

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

PE09c ik me slechts concentreer op de activiteiten binnen mijn eigen rayon

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

PE03i ik duidelijkheid krijg over wat van mij wordt verlangt

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

PE11c de rooster-privileges komen te vervallen

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Ik vind het belangrijk dat…

PE04c mijn roosterwensen buiten het capaciteitsmanagement systeem worden bijgehouden

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

PE11i ik mijn opgebouwde rooster-privileges behoud

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

PE06c ik afstand van mijn chef kan nemen

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

PE07i ik in teamverband kan werken

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Ik vind het belangrijk dat…

PE08i mijn vaardigheden worden vastgelegd

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

PE03c wat er van mij wordt verlangd in het vage blijft

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

PE12c de tijdsdruk op het werk wordt vergroot

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

PE09i de samenhang tussen activiteiten in mijn district wordt vergroot

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Ik vind het belangrijk dat…

PE05c ik het contact met mijn chef minimaal kan houden

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

PE01c de vrijheid voor eigen initiatief nihil is

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

PE10i ik in actie kan zijn

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

PE12i ik minder tijdsdruk op het werk heb

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Ik vind het belangrijk dat…

PE04i ik mijn roosterwensen kan vastleggen

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

PE05i ik door mijn chef gehoord word

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

PE07c ik alleen kan werken

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

PE02c ik alleen globale informatie over werkroosters heb

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Ik vind het belangrijk dat…

PE08c mijn vaardigheden buiten het capaciteitsmanagement systeem worden bijgehouden

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

PE01i ik vrijheid heb voor eigen initiatief

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

PE02i ik duidelijkheid krijg over de werkroosters

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

PE10c ik op het bureau me met administratief werk bezig houd

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

De volgende stellingen gaan over de eigenschappen die mogelijk in het capaciteitsmanagement systeem opgenomen worden en in hoeverre u het eens bent met deze toekomstige eigenschappen van het systeem.
U geeft uw mening door één van de getallen 0 t/m 5 te omcirkelen.

Wensen voor een capaciteitsmanagement systeem

Ik vind het belangrijk dat in een toekomstige werksituatie…

RQ03c werkactiviteiten alleen tijdens de briefing worden verantwoord (mondeling)

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

RQ11c mijn gevolgde opleidingen buiten het capaciteitsmanagement systeem worden gehouden

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

RQ01i alle werkactiviteiten vooraf worden ingeroosterd in het capaciteitsmanagement systeem

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

RQ02i het resultaat van activiteiten wordt ingevoerd in het capaciteitsmanagement systeem

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Ik vind het belangrijk dat in een toekomstige werksituatie…

RQ04c globale werkplanningen worden gemaakt

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

RQ02c het resultaat van mijn werk alleen in geschreven dagrapporten staat (buiten het systeem gelaten)

```
helemaal    oneens    enigszins    enigszins   eens      helemaal
oneens                oneens       eens                  eens
    0          1          2          3          4          5
```

RQ01c alleen diensttijden vooraf worden ingeroosterd in het capaciteitsmanagement systeem

```
helemaal    oneens    enigszins    enigszins   eens      helemaal
oneens                oneens       eens                  eens
    0          1          2          3          4          5
```

RQ06c planningen op maandbasis worden gemaakt

```
helemaal    oneens    enigszins    enigszins   eens      helemaal
oneens                oneens       eens                  eens
    0          1          2          3          4          5
```

Ik vind het belangrijk dat in een toekomstige werksituatie…

RQ04i gedetailleerde werkplanningen worden gemaakt in het capaciteitsmanagement systeem

```
helemaal    oneens    enigszins    enigszins   eens      helemaal
oneens                oneens       eens                  eens
    0          1          2          3          4          5
```

RQ03i werkactiviteiten in het capaciteitsmanagement systeem worden ingevoerd

```
helemaal    oneens    enigszins    enigszins   eens      helemaal
oneens                oneens       eens                  eens
    0          1          2          3          4          5
```

RQ08c werkplanningen worden gemaakt op basis van de hoeveelheid aanwezig personeel

```
helemaal    oneens    enigszins    enigszins   eens      helemaal
oneens                oneens       eens                  eens
    0          1          2          3          4          5
```

RQ13c roosters continu kunnen veranderen

```
helemaal    oneens    enigszins    enigszins   eens      helemaal
oneens                oneens       eens                  eens
    0          1          2          3          4          5
```

Ik vind het belangrijk dat in een toekomstige werksituatie…

RQ06i planningen op jaarbasis worden gemaakt

```
helemaal    oneens    enigszins    enigszins   eens      helemaal
oneens                oneens       eens                  eens
    0          1          2          3          4          5
```

RQ07c vakantieplanning flexibel wordt ingevuld

```
helemaal    oneens    enigszins    enigszins   eens      helemaal
oneens                oneens       eens                  eens
    0          1          2          3          4          5
```

RQ08i werkplanning op basis van vaardigheden van het personeel wordt gedaan

```
helemaal    oneens    enigszins    enigszins   eens      helemaal
oneens                oneens       eens                  eens
    0          1          2          3          4          5
```

RQ09i weinig tijd aan het urenverantwoording systeem besteed hoeft te worden

```
helemaal    oneens    enigszins    enigszins   eens      helemaal
oneens                oneens       eens                  eens
    0          1          2          3          4          5
```

Ik vind het belangrijk dat in een toekomstige werksituatie…

RQ07i vakantieplanningen volgens vaste afspraken worden gemaakt

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

RQ09c de nodige tijd aan het urenverantwoording systeem besteed wordt

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

RQ11i gevolgde opleidingen worden geregistreerd in het capaciteitsmanagement systeem

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

RQ12i roosterwensen worden vastgelegd in het capaciteitsmanagement systeem

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Ik vind het belangrijk dat in een toekomstige werksituatie…

RQ14c planningen worden gemaakt op basis van aanwezig personeel

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

RQ12c roosterwensen alleen mondeling worden toegezegd

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

RQ13i roosters 48 uur voor aanvang definitief vast worden gelegd

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

RQ14i planningen worden gemaakt op basis van verwachte drukte

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

De volgende stellingen zijn voorspellingen over de voors en tegens van een toekomstig capaciteitsmanagement systeem en in hoeverre u het eens bent met deze toekomstverwachtingen.
U geeft uw mening door één van de getallen 0 t/m 5 te omcirkelen.

Toekomstverwachtingen over een nieuw capaciteitsmanagement systeem…

PEppp1 Als mijn roosterwensen in het capaciteitsmanagement systeem worden vastgelegd, blijven mijn privileges bestaan.

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

PEnnn1 Als roosterwensen mondeling worden toegekend, verlies ik mijn privileges

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

PEppn2 Een globale werkplanning voorkomt dat ik alleen orders opvolg.

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

PEnnn3 Als roosters continu veranderen, verhoogt dat de tijdsdruk op het werk

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Toekomstverwachtingen over een nieuw capaciteitsmanagement systeem…

PEpnn3 Als roosters 48 uur voor aanvang definitief vaststaan, wordt de tijdsdruk op het werk groter

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

PEnpp3 Als roosters continu veranderen, vermindert de tijdsdruk op het werk

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

PEpnp1 Als mijn roosterwensen in CMS worden vastgelegd, hindert dat het voortbestaan van mijn privileges

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

PEpnp2 Een globale werkplanning beperkt mij in het zelf invullen van mijn werk

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Toekomstverwachtingen over een nieuw capaciteitsmanagement systeem…

PEpnp3 Als roosters 48 uur voor aanvang definitief vaststaan, verhindert dat de verlaging van de tijdsdruk

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

PEpnn1 Als mijn roosterwensen in het capaciteitsmanagement systeem worden vastgelegd, verlies ik mijn privileges

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

PEnpp1 Als roosterwensen mondeling worden toegekend, behoud ik mijn privileges

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

PEppp2 Een globale werkplanning laat ruimte voor m’n eigen invulling van mijn werk

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Toekomstverwachtingen over een nieuw capaciteitsmanagement systeem…

PEnnn2 Een gedetailleerde werkplanning heeft tot gevolg dat ik alleen orders opvolg

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens

```
    0          1          2          3          4          5
```

PEppp3 Roosters die 48 uur voor aanvang definitief vaststaan, verlagen de tijdsdruk op het werk

```
helemaal    oneens    enigszins    enigszins   eens      helemaal
oneens                oneens       eens                  eens
    0          1          2          3          4          5
```

PEnpp2 Een gedetailleerde werkplanning laat ruimte voor m'n eigen invulling van mijn werk

```
helemaal    oneens    enigszins    enigszins   eens      helemaal
oneens                oneens       eens                  eens
    0          1          2          3          4          5
```

PEnpn3 Als roosters continu veranderen, vermindert dat de tijdsdruk op het werk

```
helemaal    oneens    enigszins    enigszins   eens      helemaal
oneens                oneens       eens                  eens
    0          1          2          3          4          5
```

Toekomstverwachtingen over een nieuw capaciteitsmanagement systeem…

PEppn1 Als mijn roosterwensen in het capaciteitsmanagement systeem worden vastgelegd, voorkomt dat het verlies van mijn privileges

```
helemaal    oneens    enigszins    enigszins   eens      helemaal
oneens                oneens       eens                  eens
    0          1          2          3          4          5
```

PEnpn1 Als roosterwensen mondeling worden toegekend, behoedt mij dat voor het verliezen van mijn privileges

```
helemaal    oneens    enigszins    enigszins   eens      helemaal
oneens                oneens       eens                  eens
    0          1          2          3          4          5
```

PEnpn2 Een gedetailleerde werkplanning voorkomt dat ik alleen orders opvolg

```
helemaal    oneens    enigszins    enigszins   eens      helemaal
oneens                oneens       eens                  eens
    0          1          2          3          4          5
```

PEppn3 Als roosters 48 uur voor aanvang definitief vaststaan, voorkomt dat dat de tijdsdruk groter wordt

```
helemaal    oneens    enigszins    enigszins   eens      helemaal
oneens                oneens       eens                  eens
    0          1          2          3          4          5
```

Toekomstverwachtingen over een nieuw capaciteitsmanagement systeem…

PEnnp1 Als roosterwensen mondeling worden toegekend, is dat een obstakel voor het behouden van mijn privileges

```
helemaal    oneens    enigszins    enigszins   eens      helemaal
oneens                oneens       eens                  eens
    0          1          2          3          4          5
```

PEnnp2 Een gedetailleerde werkplanning beperkt mij in de eigen invulling van mijn werk

```
helemaal    oneens    enigszins    enigszins   eens      helemaal
oneens                oneens       eens                  eens
    0          1          2          3          4          5
```

PEnnp3 Als roosters continu veranderen, staat dat verlaging van de tijdsdruk in de weg

```
helemaal    oneens    enigszins    enigszins   eens      helemaal
oneens                oneens       eens                  eens
    0          1          2          3          4          5
```

PEpnn2 Een globale werkplanning heeft tot gevolg dat ik alleen orders opvolg

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Open vragen

D01 Leeftijd:

D02 Geslacht:

D03 Aantal jaren werkzaam bij de politie:

D04 Naam functie:

D05 Aantal jaren werkzaam in huidige functie:

D06 In het hoeveelste jaar van uw opleiding bent u:

D07 Regio:

Geef een korte beschrijving van wat u leuk vindt aan uw werk:

Geef een korte beschrijving van wat u niet leuk vindt van uw werk:

CMS RE*quest*, version P_a

Vragenlijst Capaciteit Management Politie: Groep P_a

De volgende stellingen gaan over de doelen die u persoonlijk in uw werk heeft.
U geeft uw mening door één van de getallen 0 t/m 5 te omcirkelen.

Persoonlijke doelen bij het werk
Ik vind het belangrijk dat…

PA12i ik een goede werkrelatie met mijn chef heb

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

PA01i ik mijn chef van dienst kan zijn

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

PA12c het contact met mijn chef minimaal is

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

PA08c het capaciteitsmanagement systeem de nalatigheid van mijn collega's zichtbaar maakt

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Ik vind het belangrijk dat…

PA06c ik mijn hele werkdag op het bureau doorbreng

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

PA04c het capaciteitsmanagement systeem de zwakke plekken van mijn collega's zichtbaar maakt

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

PA08i de inspanningen van mijn collega's zichtbaar worden in het capaciteitsmanagement systeem

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

PA09i ik rekening houd met de roosterwensen van collega's

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Ik vind het belangrijk dat…

PA11i ik meer betrokken word bij mijn collega's

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

PA10c collega’s hun eigen klus klaren

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

PA05i ik de vrijheid heb om spontaan mensen te helpen

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

PA01c mijn chef minder contact met mij heeft

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Ik vind het belangrijk dat…

PA02c we de dienstverlening aan de burgers kunnen verminderen

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

PA04i de kwaliteiten van mijn collega’s tot hun recht komen

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

PA03c ik me alleen op mijn eigen werkzaamheden hoef te richten

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

PA10i dubbel werk wordt voorkomen

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Ik vind het belangrijk dat…

PA06i ik mijn werkdag op straat kan doorbrengen

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

PA07i ik samenwerk met mijn collega’s

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

PA05c ik alleen mensen hoef te helpen als daarvoor tijd gepland staat

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

PA09c de roosterwensen van mijn collega’s worden genegeerd

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Ik vind het belangrijk dat…

PA11c ik voor mezelf mag werken

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

PA02i ik de dienstverlening aan de burgers kan verbeteren

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

PA03i ik tijd overhoud om mijn collega's te helpen

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

PA07c ieder voor zich werkt

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

De volgende stellingen gaan over de eigenschappen die mogelijk in het capaciteits-management systeem opgenomen worden en in hoeverre u het eens bent met deze toekomstige eigenschappen van het systeem.
U geeft uw mening door één van de getallen 0 t/m 5 te omcirkelen.

Wensen voor een capaciteitsmanagement systeem

Ik vind het belangrijk dat in een toekomstige werksituatie…

RQ03c werkactiviteiten alleen tijdens de briefing worden verantwoord (mondeling)

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

RQ11c mijn gevolgde opleidingen buiten het capaciteitsmanagement systeem worden gehouden

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

RQ01i alle werkactiviteiten vooraf worden ingeroosterd in het capaciteitsmanagement systeem

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

RQ02i het resultaat van activiteiten wordt ingevoerd in het capaciteitsmanagement systeem

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Ik vind het belangrijk dat in een toekomstige werksituatie…

RQ04c globale werkplanningen worden gemaakt

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

RQ02c het resultaat van mijn werk alleen in geschreven dagrapporten staat (buiten het systeem gelaten)

```
helemaal    oneens    enigszins    enigszins   eens      helemaal
oneens                oneens       eens                  eens
    0          1          2           3          4          5
```

RQ01c alleen diensttijden vooraf worden ingeroosterd in het capaciteitsmanagement systeem

```
helemaal    oneens    enigszins    enigszins   eens      helemaal
oneens                oneens       eens                  eens
    0          1          2           3          4          5
```

RQ06c planningen op maandbasis worden gemaakt

```
helemaal    oneens    enigszins    enigszins   eens      helemaal
oneens                oneens       eens                  eens
    0          1          2           3          4          5
```

Ik vind het belangrijk dat in een toekomstige werksituatie…

RQ04i gedetailleerde werkplanningen worden gemaakt in het capaciteitsmanagement systeem

```
helemaal    oneens    enigszins    enigszins   eens      helemaal
oneens                oneens       eens                  eens
    0          1          2           3          4          5
```

RQ03i werkactiviteiten in het capaciteitsmanagement systeem worden ingevoerd

```
helemaal    oneens    enigszins    enigszins   eens      helemaal
oneens                oneens       eens                  eens
    0          1          2           3          4          5
```

RQ08c werkplanningen worden gemaakt op basis van de hoeveelheid aanwezig personeel

```
helemaal    oneens    enigszins    enigszins   eens      helemaal
oneens                oneens       eens                  eens
    0          1          2           3          4          5
```

RQ13c roosters continu kunnen veranderen

```
helemaal    oneens    enigszins    enigszins   eens      helemaal
oneens                oneens       eens                  eens
    0          1          2           3          4          5
```

Ik vind het belangrijk dat in een toekomstige werksituatie…

RQ06i planningen op jaarbasis worden gemaakt

```
helemaal    oneens    enigszins    enigszins   eens      helemaal
oneens                oneens       eens                  eens
    0          1          2           3          4          5
```

RQ07c vakantieplanning flexibel wordt ingevuld

```
helemaal    oneens    enigszins    enigszins   eens      helemaal
oneens                oneens       eens                  eens
    0          1          2           3          4          5
```

RQ08i werkplanning op basis van vaardigheden van het personeel wordt gedaan

```
helemaal    oneens    enigszins    enigszins   eens      helemaal
oneens                oneens       eens                  eens
    0          1          2           3          4          5
```

RQ09i weinig tijd aan het urenverantwoording systeem besteed hoeft te worden

```
helemaal    oneens    enigszins    enigszins   eens      helemaal
oneens                oneens       eens                  eens
    0          1          2           3          4          5
```

Ik vind het belangrijk dat in een toekomstige werksituatie…

RQ07i vakantieplanningen volgens vaste afspraken worden gemaakt

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

RQ09c de nodige tijd aan het urenverantwoording systeem besteed wordt

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

RQ11i gevolgde opleidingen worden geregistreerd in het capaciteitsmanagement systeem

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

RQ12i roosterwensen worden vastgelegd in het capaciteitsmanagement systeem

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Ik vind het belangrijk dat in een toekomstige werksituatie…

RQ14c planningen worden gemaakt op basis van aanwezig personeel

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

RQ12c roosterwensen alleen mondeling worden toegezegd

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

RQ13i roosters 48 uur voor aanvang definitief vast worden gelegd

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

RQ14i planningen worden gemaakt op basis van verwachte drukte

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

De volgende stellingen zijn voorspellingen over de voors en tegens van een toekomstig capaciteitsmanagement systeem en in hoeverre u het eens bent met deze toekomstverwachtingen.
U geeft uw mening door één van de getallen 0 t/m 5 te omcirkelen.

Toekomstverwachtingen over een nieuw capaciteitsmanagement systeem…

PAnpp3 Door mijn werkactiviteiten mondeling te verantwoorden ben ik mijn chef van dienst

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

PAppp2 Alle werkactiviteiten vooraf inroosteren komt de dienstverlening aan de burger ten goede

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

PAnnn3 Door mijn werkactiviteiten mondeling te verantwoorden werk ik mijn chef tegen

helemaal oneens	oneens	enigszins	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

PAppn1 Dat mijn rooster 48 uur van tevoren vaststaat voorkomt dat ik alleen voor mezelf werk

helemaal oneens	oneens	enigszins	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Toekomstverwachtingen over een nieuw capaciteitsmanagement systeem…

PAppn2 Door het vooraf inroosteren van alle werkactiviteiten voorkom je dat de burger aan zijn lot wordt overgelaten

helemaal oneens	oneens	enigszins	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

PAnnp2 Alleen diensttijden inroosteren heeft negatieve gevolgen voor de dienstverlening aan de burger

helemaal oneens	oneens	enigszins	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

PApnp1 Als mijn rooster 48 uur van tevoren vaststaat, verslechtert dat mijn samenwerking met de collega's

helemaal oneens	oneens	enigszins	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

PApnp3 Door mijn werkactiviteiten te verantwoorden in het capaciteitsmanagement systeem ben ik mijn chef slecht van dienst

helemaal oneens	oneens	enigszins	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Toekomstverwachtingen over een nieuw capaciteitsmanagement systeem…

PApnp2 Alle werkactiviteiten vooraf inroosteren heeft negatieve gevolgen voor de dienstverlening aan de burger

helemaal oneens	oneens	enigszins	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

PAnpn3 Dat ik mijn werkactiviteiten mondeling verantwoord voorkomt dat ik mijn chef tegenwerk

helemaal oneens	oneens	enigszins	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

PApnn2 Door het vooraf inroosteren van alle werkactiviteiten wordt de burger aan zijn lot overgelaten

helemaal oneens	oneens	enigszins	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

PAppn3 Doordat ik mijn werkactiviteiten verantwoord in het capaciteitsmanagement systeem voorkom ik dat ik mijn chef tot last ben

helemaal oneens	oneens	enigszins	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Toekomstverwachtingen over een nieuw capaciteitsmanagement systeem…

PAnpp1 Als de werkroosters continu veranderen bevordert dat mijn samenwerking met de collega's

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

PApnn3 Werkactiviteiten verantwoorden in het capaciteitsmanagement systeem maakt dat ik mijn chef meer tot last ben

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

PAnnn2 Door alleen diensttijden in te roosteren wordt de burger aan zijn lot overgelaten

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

PAnpn1 Als de werkroosters continu veranderen voorkomt dat dat ik alleen voor mezelf werk

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Toekomstverwachtingen over een nieuw capaciteitsmanagement systeem…

PAppp3 Door mijn werkactiviteiten te verantwoorden in het capaciteitsmanagement systeem ben ik mijn chef goed van dienst

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

PAppp1 Dat mijn rooster 48 uur van tevoren vaststaat, bevordert mijn samenwerking met de collega's

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

PAnpp2 Alleen werktijden vooraf inroosteren, komt de dienstverlening aan de burger ten goede

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

PAnnp1 Dat de werkroosters continu veranderen verslechtert mijn samenwerking met de collega's

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Toekomstverwachtingen over een nieuw capaciteitsmanagement systeem…

PAnnp3 Door mijn werkactiviteiten mondeling te verantwoorden ben ik mijn chef slecht van dienst

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

PAnnn1 Doordat de werkroosters continu veranderen werk ik alleen voor mezelf

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

PApnn1 Omdat mijn rooster al 48 uur van tevoren vaststaat werk ik alleen voor mezelf

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

PAnpn2 Met alleen inroosteren van de diensttijden vermijd je dat de burger aan zijn lot wordt overgelaten

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Open vragen

D01 Leeftijd:

D02 Geslacht:

D03 Aantal jaren werkzaam bij de politie:

D04 Naam functie:

D05 Aantal jaren werkzaam in huidige functie:

D06 In het hoeveelste jaar van uw opleiding bent u:

D07 Regio:

Geef een korte beschrijving van wat u leuk vindt aan uw werk:

Geef een korte beschrijving van wat u niet leuk vindt van uw werk:

CMS RE*quest*, version B_e

Vragenlijst Capaciteit Management Politie: Groep B_e

De volgende stellingen gaan over de doelen van uw korps en in hoeverre u het eens bent met deze doelen.
U geeft uw mening door één van de getallen 0 t/m 5 te omcirkelen.

Doelen van het korps
Ik vind het belangrijk dat mijn korps…

BE08i Het ziekteverzuim onder het personeel vermindert

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

BE11c De activiteiten van de BPZ afhankelijk laat zijn van de situatie

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

BE10i Professioneel overkomt op de buitenwereld

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

BE09c Meer bemoeienis van BZK toelaat

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Ik vind het belangrijk dat mijn korps…

BE06c Het personeel van de BPZ vrijlaat in hun werkzaamheden

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

BE02i Geld bespaart op personeelsinzet

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

BE05c Tegen de politiek ingaat

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

BE04i Rust op de werkvloer houdt

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Ik vind het belangrijk dat mijn korps…

BE01i Het prestatiecontract haalt

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

BE10c Amateuristisch overkomt op de buitenwereld

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

BE05i De politiek tegemoet komt

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

BE04c Onrust op de werkvloer creëert

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Ik vind het belangrijk dat mijn korps…

BE12i Samenwerking tussen de verschillende districten bevordert

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

BE03c De werkroosters continu wijzigt

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

BE06i De werkzaamheden van de BPZ goed aanstuurt

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

BE02c Geld toelegt op personeelsinzet

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Ik vind het belangrijk dat mijn korps…

BE07i De activiteiten van de BPZ controleert

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

BE08c Het ziekteverzuim onder het personeel laat voor wat het is

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

BE09i De bemoeienis van BZK beperkt kan houden

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

BE12c Samenwerking tussen districten vermijdt

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Ik vind het belangrijk dat mijn korps…

BE03i De werkroosters op tijd plant

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens

```
    0          1          2          3          4          5
```

BE01c Het prestatiecontract erbij laat zitten

```
helemaal    oneens    enigszins    enigszins    eens      helemaal
oneens                oneens       eens                   eens
    0          1          2          3          4          5
```

BE11i De activiteiten van de BPZ plant

```
helemaal    oneens    enigszins    enigszins    eens      helemaal
oneens                oneens       eens                   eens
    0          1          2          3          4          5
```

BE07c De activiteiten van de BPZ controle-vrij laat

```
helemaal    oneens    enigszins    enigszins    eens      helemaal
oneens                oneens       eens                   eens
    0          1          2          3          4          5
```

De volgende stellingen gaan over de eigenschappen die mogelijk in het capaciteitsmanagement systeem opgenomen worden en in hoeverre u het eens bent met deze toekomstige eigenschappen van het systeem.
U geeft uw mening door één van de getallen 0 t/m 5 te omcirkelen.

Wensen voor een capaciteitsmanagement systeem

Ik vind het belangrijk dat in een toekomstige werksituatie…

RQ03c werkactiviteiten alleen tijdens de briefing worden verantwoord (mondeling)

```
helemaal    oneens    enigszins    enigszins    eens      helemaal
oneens                oneens       eens                   eens
    0          1          2          3          4          5
```

RQ11c mijn gevolgde opleidingen buiten het capaciteitsmanagement systeem worden gehouden

```
helemaal    oneens    enigszins    enigszins    eens      helemaal
oneens                oneens       eens                   eens
    0          1          2          3          4          5
```

RQ01i alle werkactiviteiten vooraf worden ingeroosterd in het capaciteitsmanagement systeem

```
helemaal    oneens    enigszins    enigszins    eens      helemaal
oneens                oneens       eens                   eens
    0          1          2          3          4          5
```

RQ02i het resultaat van activiteiten wordt ingevoerd in het capaciteitsmanagement systeem

```
helemaal    oneens    enigszins    enigszins    eens      helemaal
oneens                oneens       eens                   eens
    0          1          2          3          4          5
```

Ik vind het belangrijk dat in een toekomstige werksituatie…

RQ04c globale werkplanningen worden gemaakt

```
helemaal    oneens    enigszins    enigszins    eens      helemaal
oneens                oneens       eens                   eens
    0          1          2          3          4          5
```

RQ02c het resultaat van mijn werk alleen in geschreven dagrapporten staat (buiten het systeem gelaten)

```
helemaal    oneens    enigszins    enigszins    eens      helemaal
oneens                oneens       eens                   eens
    0          1          2          3          4          5
```

RQ01c alleen diensttijden vooraf worden ingeroosterd in het capaciteitsmanagement systeem

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

RQ06c planningen op maandbasis worden gemaakt

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Ik vind het belangrijk dat in een toekomstige werksituatie…

RQ04i gedetailleerde werkplanningen worden gemaakt in het capaciteitsmanagement systeem

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

RQ03i werkactiviteiten in het capaciteitsmanagement systeem worden ingevoerd

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

RQ08c werkplanningen worden gemaakt op basis van de hoeveelheid aanwezig personeel

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

RQ13c roosters continu kunnen veranderen

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Ik vind het belangrijk dat in een toekomstige werksituatie…

RQ06i planningen op jaarbasis worden gemaakt

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

RQ07c vakantieplanning flexibel wordt ingevuld

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

RQ08i werkplanning op basis van vaardigheden van het personeel wordt gedaan

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

RQ09i weinig tijd aan het urenverantwoording systeem besteed hoeft te worden

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Ik vind het belangrijk dat in een toekomstige werksituatie…

RQ07i vakantieplanningen volgens vaste afspraken worden gemaakt

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

RQ09c de nodige tijd aan het urenverantwoording systeem besteed wordt

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

RQ11i gevolgde opleidingen worden geregistreerd in het capaciteitsmanagement systeem

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

RQ12i roosterwensen worden vastgelegd in het capaciteitsmanagement systeem

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Ik vind het belangrijk dat in een toekomstige werksituatie…

RQ14c planningen worden gemaakt op basis van aanwezig personeel

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

RQ12c roosterwensen alleen mondeling worden toegezegd

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

RQ13i roosters 48 uur voor aanvang definitief vast worden gelegd

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

RQ14i planningen worden gemaakt op basis van verwachte drukte

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

De volgende stellingen zijn voorspellingen over de voors en tegens van een toekomstig capaciteitsmanagement systeem en in hoeverre u het eens bent met deze toekomstverwachtingen.
U geeft uw mening door één van de getallen 0 t/m 5 te omcirkelen.

Toekomstverwachtingen over een nieuw capaciteitsmanagement systeem…

BEnpp1 Planning op basis van aanwezig personeel bespaart geld op personeelsinzet

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

BEpnp2 Een gedetailleerde werkplanning belet ons het prestatiecontract te halen

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

BEnpp3 Door resultaten alleen vast te leggen in dagrapporten (buiten het systeem) komen we professioneel over

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

BEppp2 Een gedetailleerde werkplanning helpt het prestatiecontract te halen

helemaal	oneens	enigszins	enigszins	eens	helemaal

```
oneens              oneens       eens                eens
    0         1         2         3         4         5
```

Toekomstverwachtingen over een nieuw capaciteitsmanagement systeem…

BEppn2 Een gedetailleerde werkplanning voorkomt dat we het prestatiecontract verzaken

```
helemaal    oneens    enigszins    enigszins    eens      helemaal
oneens                oneens       eens                   eens
    0         1         2         3         4         5
```

BEppn3 Door resultaten in het capaciteitsmanagement systeem vast te leggen vermijden we amateurisme
binnen de BPZ

```
helemaal    oneens    enigszins    enigszins    eens      helemaal
oneens                oneens       eens                   eens
    0         1         2         3         4         5
```

BEppp1 Planning op basis van de verwachte drukte bespaart geld op personeelsinzet

```
helemaal    oneens    enigszins    enigszins    eens      helemaal
oneens                oneens       eens                   eens
    0         1         2         3         4         5
```

BEnpp2 Een globale werkplanning helpt het prestatiecontract te halen

```
helemaal    oneens    enigszins    enigszins    eens      helemaal
oneens                oneens       eens                   eens
    0         1         2         3         4         5
```

Toekomstverwachtingen over een nieuw capaciteitsmanagement systeem…

BEnpn3 Door resultaten vast te leggen in dagrapporten voorkomen we amateurisme

```
helemaal    oneens    enigszins    enigszins    eens      helemaal
oneens                oneens       eens                   eens
    0         1         2         3         4         5
```

BEnpn2 Een globale werkplanning voorkomt dat we het prestatiecontract verzuimen

```
helemaal    oneens    enigszins    enigszins    eens      helemaal
oneens                oneens       eens                   eens
    0         1         2         3         4         5
```

BEpnp1 Planning op basis van de verwachte drukte remt een besparing op personeelsinzet

```
helemaal    oneens    enigszins    enigszins    eens      helemaal
oneens                oneens       eens                   eens
    0         1         2         3         4         5
```

BEnnn3 Door resultaten vast te leggen in dagrapporten komen we amateuristisch over

```
helemaal    oneens    enigszins    enigszins    eens      helemaal
oneens                oneens       eens                   eens
    0         1         2         3         4         5
```

Toekomstverwachtingen over een nieuw capaciteitsmanagement systeem…

BEnnp1 Planning op basis van aanwezig personeel staat een besparing op personeelskosten in de weg

```
helemaal    oneens    enigszins    enigszins    eens      helemaal
oneens                oneens       eens                   eens
    0         1         2         3         4         5
```

BEnnn2 Door de werkplanning globaal te houden zal het met het prestatiecontract mislopen

```
helemaal    oneens    enigszins    enigszins    eens      helemaal
oneens                oneens       eens                   eens
    0         1         2         3         4         5
```

BEpnp3 Het vastleggen van resultaten in het capaciteitsmanagement systeem schaadt ons imago als professionals

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

BEnnp2 Een globale werkplanning belet ons het prestatiecontract te halen

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Toekomstverwachtingen over een nieuw capaciteitsmanagement systeem…

BEpnn1 Planning op basis van de verwachte drukte verhoogt de kosten van personeelsinzet

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

BEnnp3 Het vastleggen van resultaten in dagrapporten schaadt ons imago als professionals

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

BEnnn1 Planning op basis van aanwezig personeel leidt tot hoge kosten voor personeelsinzet

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

BEpnn3 Door resultaten vast te leggen in het capaciteitsmanagement systeem komen we amateuristisch over

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Toekomstverwachtingen over een nieuw capaciteitsmanagement systeem…

BEnpn1 Planning op basis van aanwezig personeel verlaagt de kosten van personeelsinzet

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

BEppp3 Door resultaten in het capaciteitsmanagement systeem vast te leggen komen we professioneel over

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

BEppn1 Planning op basis van de verwachte drukte zorgt voor een stijging in de personeelskosten

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

BEpnn2 Een gedetailleerde werkplanning maakt dat we het prestatiecontract gaan verwaarlozen

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Open vragen

D01	Leeftijd:
D02	Geslacht:
D03	Aantal jaren werkzaam bij de politie:
D04	Naam functie:
D05	Aantal jaren werkzaam in huidige functie:
D06	In het hoeveelste jaar van uw opleiding bent u:
D07	Regio:

Geef een korte beschrijving van wat u leuk vindt aan uw werk:

Geef een korte beschrijving van wat u niet leuk vindt van uw werk:

CMS RE*quest*, version B_a

Vragenlijst Capaciteit Management Politie: Groep B_a

De volgende stellingen gaan over de doelen van uw korps en in hoeverre u het eens bent met deze doelen.
U geeft uw mening door één van de getallen 0 t/m 5 te omcirkelen.

Doelen van het korps
Ik vind het belangrijk dat mijn korps…

BA08i criminelen opspoort

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

BA10c de afstand tot de burgers groter maakt

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

BA01c kleine criminaliteit buiten beschouwing laat

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

BA05i goede dienstverlening aan de burger biedt

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Ik vind het belangrijk dat mijn korps…

BA11c misdaadpreventie een lage prioriteit geeft

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

BA03c het gevoel van veiligheid bij burgers buiten beschouwing laat

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

BA06i misdrijven kan oplossen

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

BA10i burgers persoonlijk te woord kan staan

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Ik vind het belangrijk dat mijn korps…

BA11i burgers voorlicht over misdaadpreventie

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

BA06c het oplossen van misdrijven als een bijkomstigheid ziet

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

BA12i zichtbaar is voor de burger

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

BA07c veelplegers laat lopen

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Ik vind het belangrijk dat mijn korps…

BA01i kleine criminaliteit bestrijdt

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

BA04i de openbare orde handhaaft

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

BA02c terrorismebestrijding aan de AIVD overlaat

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

BA04c weinig doet aan handhaving van de openbare orde

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Ik vind het belangrijk dat mijn korps…

BA05c minimale dienstverlening aan de burger biedt

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

BA07i veelplegers aanpakt

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

BA09c de verkeersveiligheid buiten beschouwing laat

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

BA03i gevoel van veiligheid bij burgers verbetert

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Ik vind het belangrijk dat mijn korps…

BA02i terrorisme kan bestrijden

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

BA12c onzichtbaar is voor de burger

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

BA08c criminelen laat lopen

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

BA09i de verkeersveiligheid vergroot

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

De volgende stellingen gaan over de eigenschappen die mogelijk in het capaciteits-management systeem opgenomen worden en in hoeverre u het eens bent met deze toekomstige eigenschappen van het systeem.
U geeft uw mening door één van de getallen 0 t/m 5 te omcirkelen.

Wensen voor een capaciteitsmanagement systeem

Ik vind het belangrijk dat in een toekomstige werksituatie…

RQ03c werkactiviteiten alleen tijdens de briefing worden verantwoord (mondeling)

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

RQ11c mijn gevolgde opleidingen buiten het capaciteitsmanagement systeem worden gehouden

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

RQ01i alle werkactiviteiten vooraf worden ingeroosterd in het capaciteitsmanagement systeem

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

RQ02i het resultaat van activiteiten wordt ingevoerd in het capaciteitsmanagement systeem

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Ik vind het belangrijk dat in een toekomstige werksituatie…

RQ04c globale werkplanningen worden gemaakt

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

RQ02c het resultaat van mijn werk alleen in geschreven dagrapporten staat (buiten het systeem gelaten)

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

RQ01c alleen diensttijden vooraf worden ingeroosterd in het capaciteitsmanagement systeem

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

RQ06c planningen op maandbasis worden gemaakt

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Ik vind het belangrijk dat in een toekomstige werksituatie…

RQ04i gedetailleerde werkplanningen worden gemaakt in het capaciteitsmanagement systeem

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

RQ03i werkactiviteiten in het capaciteitsmanagement systeem worden ingevoerd

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

RQ08c werkplanningen worden gemaakt op basis van de hoeveelheid aanwezig personeel

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

RQ13c roosters continu kunnen veranderen

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Ik vind het belangrijk dat in een toekomstige werksituatie…

RQ06i planningen op jaarbasis worden gemaakt

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

RQ07c vakantieplanning flexibel wordt ingevuld

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

RQ08i werkplanning op basis van vaardigheden van het personeel wordt gedaan

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

RQ09i weinig tijd aan het urenverantwoording systeem besteed hoeft te worden

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Ik vind het belangrijk dat in een toekomstige werksituatie…

RQ07i vakantieplanningen volgens vaste afspraken worden gemaakt

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

RQ09c de nodige tijd aan het urenverantwoording systeem besteed wordt

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

RQ11i gevolgde opleidingen worden geregistreerd in het capaciteitsmanagement systeem

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

RQ12i roosterwensen worden vastgelegd in het capaciteitsmanagement systeem

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Ik vind het belangrijk dat in een toekomstige werksituatie…

RQ14c planningen worden gemaakt op basis van aanwezig personeel

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

RQ12c roosterwensen alleen mondeling worden toegezegd

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

RQ13i roosters 48 uur voor aanvang definitief vast worden gelegd

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

RQ14i planningen worden gemaakt op basis van verwachte drukte

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

De volgende stellingen zijn voorspellingen over de voors en tegens van een toekomstig capaciteitsmanagement systeem en in hoeverre u het eens bent met deze toekomstverwachtingen.
U geeft uw mening door één van de getallen 0 t/m 5 te omcirkelen.

Toekomstverwachtingen over een nieuw capaciteitsmanagement systeem…

BAppn2 Planning op grond van vaardigheden van het personeel voorkomt dat grootschalige optredens mislukken

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

BAnpp2 Planning op basis van beschikbaar personeel garandeert meer resultaat bij grootschalig optreden

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

BAnnp3 Dagrapportage buiten het systeem om blokkeert een snelle afhandeling van maatschappelijk dringende zaken

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

BApnn2 Planning op basis van vaardigheden van personeel leidt ertoe dat grootschalige optredens minder resultaat hebben

helemaal oneens	oneens	enigszins	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Toekomstverwachtingen over een nieuw capaciteitsmanagement systeem…

BApnp1 Planning van inzet op jaarbasis belemmert onze aansluiting op de verwachting van de burger

helemaal oneens	oneens	enigszins	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

BApnn1 Planning van inzet op jaarbasis maakt dat we minder aansluiting hebben met de verwachting van de burger

helemaal oneens	oneens	enigszins	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

BAppn1 Planning van inzet op jaarbasis voorkomt dat we aansluiting met de burger missen

helemaal oneens	oneens	enigszins	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

BAnnn2 Planning op grond van beschikbaar personeel leidt ertoe dat grootschalige optredens minder effectief zijn

helemaal oneens	oneens	enigszins	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Toekomstverwachtingen over een nieuw capaciteitsmanagement systeem…

BApnn3 Door rapportage in het capaciteitsmanagement systeem blijven maatschappelijk dringende zaken lang liggen

helemaal oneens	oneens	enigszins	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

BApnp3 Rapportage in het capaciteitsmanagement systeem stimuleert een snelle afhandeling van maatschappelijk dringende zaken

helemaal oneens	oneens	enigszins	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

BAnpn2 Planning op grond van beschikbaar personeel voorkomt dat bijzondere acties uit de hand lopen

helemaal oneens	oneens	enigszins	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

BAnpp3 Door dagrapportage buiten het systeem om worden maatschappelijk dringende zaken snel afgehandeld

helemaal oneens	oneens	enigszins	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Toekomstverwachtingen over een nieuw capaciteitsmanagement systeem…

BAnpn3 Dagrapportage buiten het systeem om verhindert dat maatschappelijk dringende zaken lang blijven liggen

helemaal oneens	oneens	enigszins	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

BAnnp1 Planning van inzet op maandbasis belemmert onze tegemoetkoming aan de verwachting van de burger

helemaal oneens	oneens	enigszins	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

BAppp1 Planning van inzet op jaarbasis maakt dat we beter aansluiten bij de verwachting van de burger

helemaal oneens	oneens	enigszins	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

BAnnn3 Door dagrapportage buiten het systeem om blijven maatschappelijk dringende zaken lang liggen

helemaal oneens	oneens	enigszins	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Toekomstverwachtingen over een nieuw capaciteitsmanagement systeem…

BAnpn1 Planning van inzet op maandbasis voorkomt dat we de aansluiting met de burger verliezen

helemaal oneens	oneens	enigszins	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

BApnp2 Planning op grond van vaardigheden van het personeel vermindert het resultaat van grootschalig optreden

helemaal oneens	oneens	enigszins	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

BAppp3 Door rapportage in het capaciteitsmanagement systeem worden maatschappelijk dringende zaken snel opgevolgd

helemaal oneens	oneens	enigszins	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

BAnpp1 Planning van inzet op maandbasis maakt dat we beter aansluiten bij de verwachting van de burger

helemaal oneens	oneens	enigszins	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Toekomstverwachtingen over een nieuw capaciteitsmanagement systeem…

BAppp2 Planning op grond van vaardigheden van personeel garandeert meer resultaat bij grootschalig optreden

helemaal oneens	oneens	enigszins	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

BAnnp2 Planning op grond van beschikbaar personeel vermindert het resultaat van grootschalig optreden

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

BAnnn1 Planning van inzet op maandbasis zorgt ervoor dat we tegemoet komen aan de verwachting van de burger

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

BAppn3 Rapportage in het capaciteitsmanagement systeem voorkomt dat maatschappelijk dringende zaken aan onze aandacht ontsnappen

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Open vragen

D01 Leeftijd:

D02 Geslacht:

D03 Aantal jaren werkzaam bij de politie:

D04 Naam functie:

D05 Aantal jaren werkzaam in huidige functie:

D06 In het hoeveelste jaar van uw opleiding bent u:

D07 Regio:

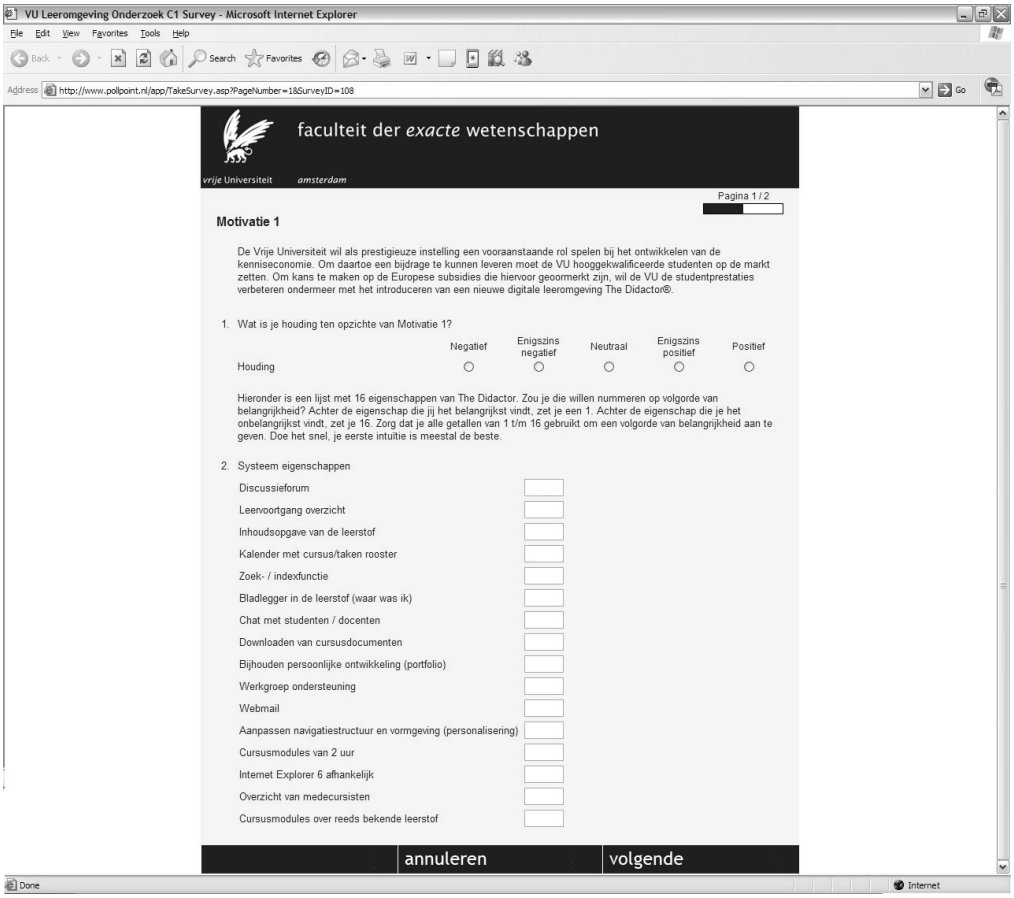
Geef een korte beschrijving van wat u leuk vindt aan uw werk:

Geef een korte beschrijving van wat u niet leuk vindt van uw werk:

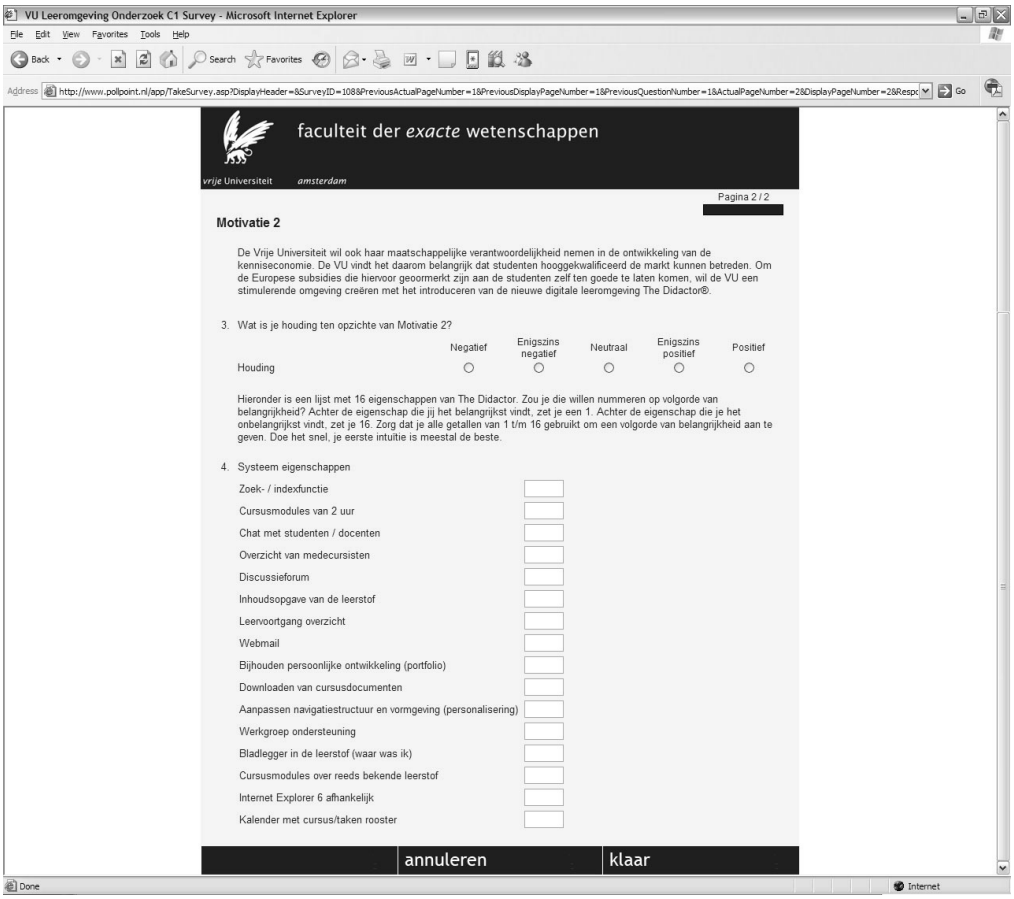
Appendix 3.2

Survey Screenshots

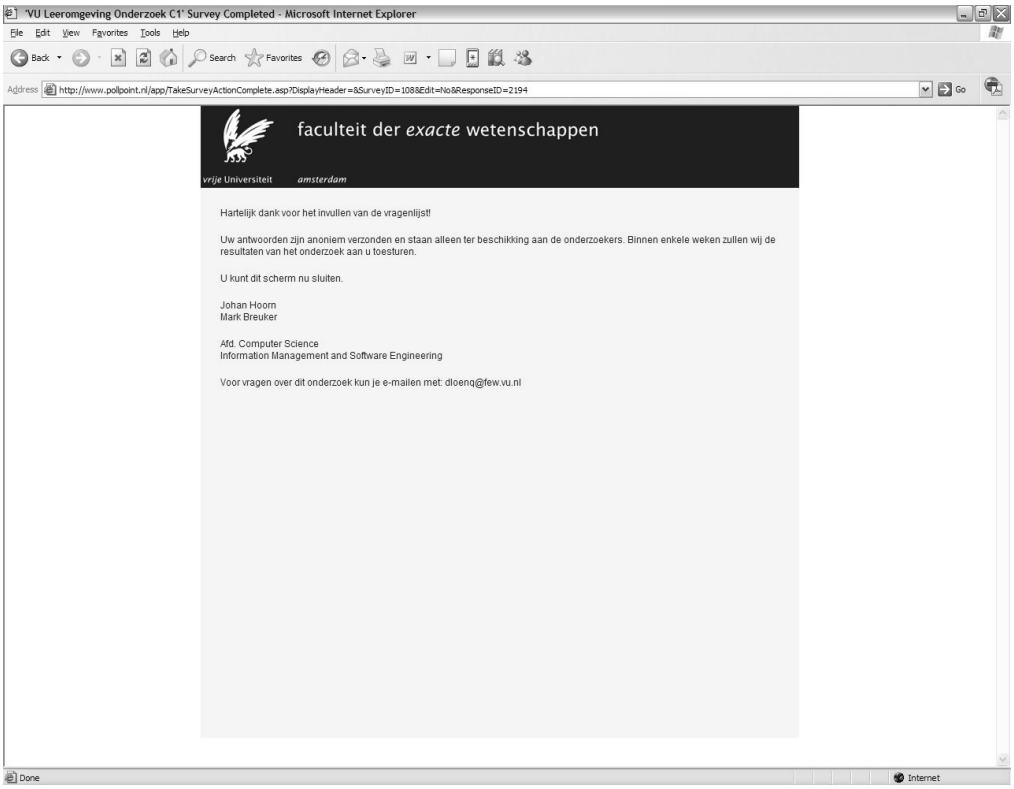
The following images are example screenshots of the online survey as it was presented to the science students.



Page 1


VU Leeromgeving Onderzoek C1 Survey - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Back Search Favorites
Address http://www.pollpoint.nl/app/TakeSurvey.asp?DisplayHeader=&SurveyID=1088&PreviousActualPageNumber=1&PreviousDisplayPageNumber=1&PreviousQuestionNumber=1&ActualPageNumber=2&DisplayPageNumber=2&Resp
Go
faculteit der exacte wetenschappen
vrije Universiteit amsterdam
Pagina 2 / 2
Motivatie 2
De Vrije Universiteit wil ook haar maatschappelijke verantwoordelijkheid nemen in de ontwikkeling van de kenniseconomie. De VU vindt het daarom belangrijk dat studenten hooggekwalificeerd de markt kunnen betreden. Om de Europese subsidies die hiervoor geoormerkt zijn aan de studenten zelf ten goede te laten komen, wil de VU een stimulerende omgeving creëren met het introduceren van de nieuwe digitale leeromgeving The Didactor®.
3. Wat is je houding ten opzichte van Motivatie 2?
Negatief
Enigszins negatief
Neutraal
Enigszins positief
Positief
Houding
Hieronder is een lijst met 16 eigenschappen van The Didactor. Zou je die willen nummeren op volgorde van belangrijkheid? Achter de eigenschap die jij het belangrijkst vindt, zet je een 1. Achter de eigenschap die je het onbelangrijkst vindt, zet je 16. Zorg dat je alle getallen van 1 t/m 16 gebruikt om een volgorde van belangrijkheid aan te geven. Doe het snel, je eerste intuitie is meestal de beste.
4. Systeem eigenschappen
Zoek- / indexfunctie
Cursusmodules van 2 uur
Chat met studenten / docenten
Overzicht van medecursisten
Discussieforum
Inhoudsopgave van de leerstof
Leervoortgang overzicht
Webmail
Bijhouden persoonlijke ontwikkeling (portfolio)
Downloaden van cursusdocumenten
Aanpassen navigatiestructuur en vormgeving (personalisering)
Werkgroep ondersteuning
Bladlegger in de leerstof (waar was ik)
Cursusmodules over reeds bekende leerstof
Internet Explorer 6 afhankelijk
Kalender met cursus/taken rooster
annuleren
klaar
Done
Internet

Page 2

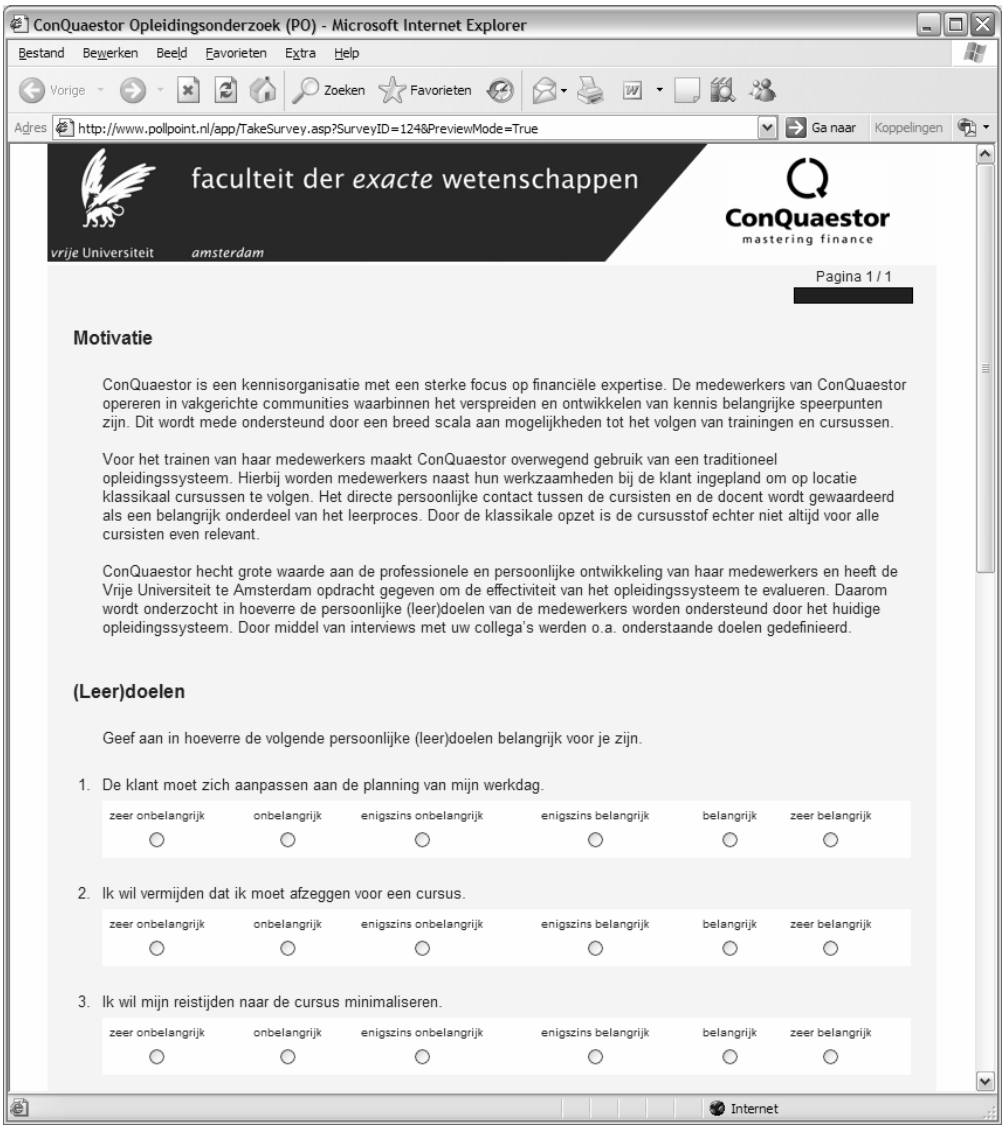

'VU Leeromgeving Onderzoek C1' Survey Completed - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Back Search Favorites
Address http://www.pollpoint.nl/app/TakeSurveyActionComplete.asp?DisplayHeader=&SurveyID=1088&Edit=No&ResponseID=2194
Go
faculteit der exacte wetenschappen
vrije Universiteit amsterdam
Hartelijk dank voor het invullen van de vragenlijst!
Uw antwoorden zijn anoniem verzonden en staan alleen ter beschikking aan de onderzoekers. Binnen enkele weken zullen wij de resultaten van het onderzoek aan u toesturen.
U kunt dit scherm nu sluiten.
Johan Hoorn
Mark Breuker
Afd. Computer Science
Information Management and Software Engineering
Voor vragen over dit onderzoek kun je e-mailen met: dloenq@few.vu.nl
Done
Internet

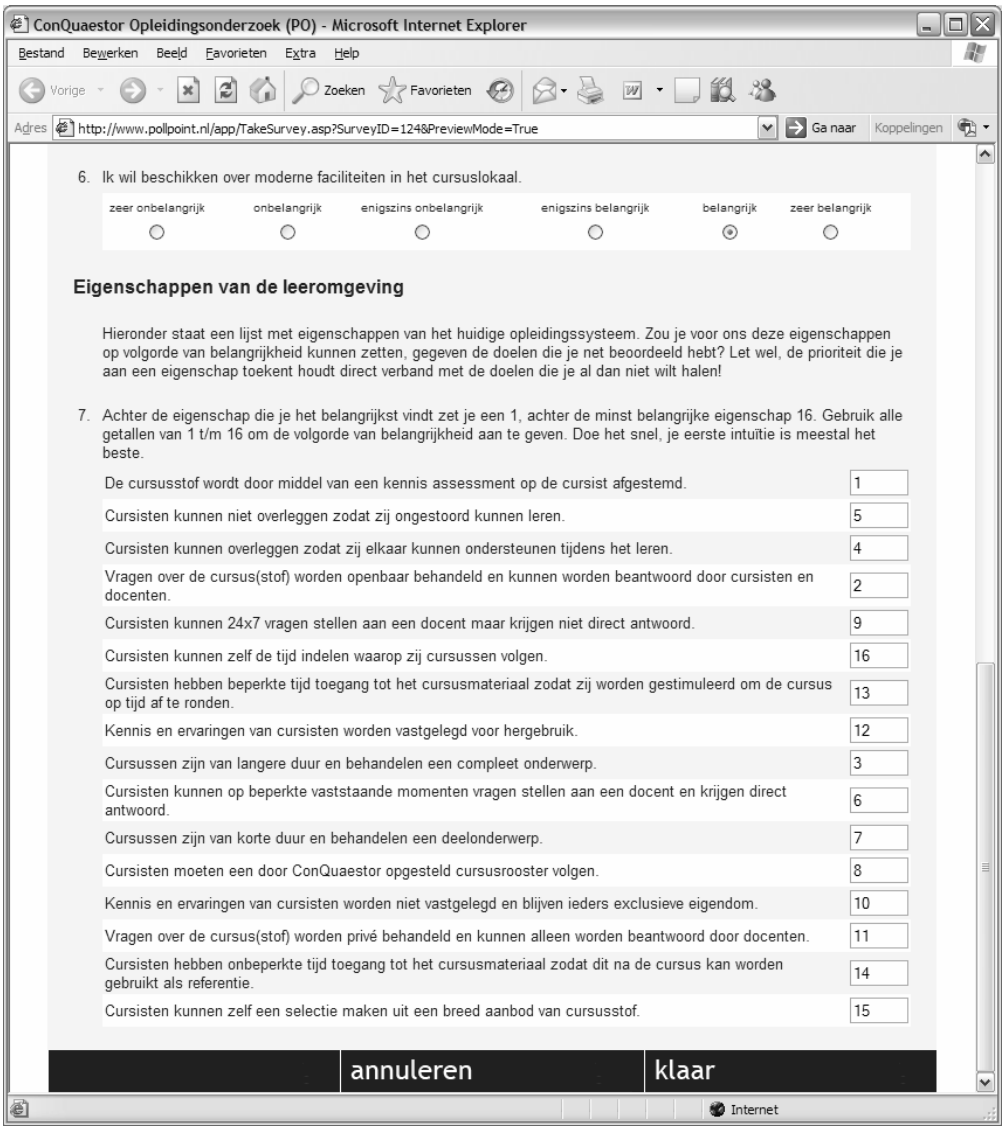
Page 3

Appendix 3.3

Survey Screenshots

The following images are example screenshots of the online survey as it was presented to the financial consultants.




ConQuaestor Opleidingsonderzoek (PO) - Microsoft Internet Explorer
Bestand Bewerken Beeld Favorieten Extra Help
Vorige Zoeken Favorieten
Adres http://www.pollpoint.nl/app/TakeSurvey.asp?SurveyID=1248&PreviewMode=True
Ga naar Koppelingen
6. Ik wil beschikken over moderne faciliteiten in het cursuslokaal.
zeer onbelangrijk
onbelangrijk
enigszins onbelangrijk
enigszins belangrijk
belangrijk
zeer belangrijk
Eigenschappen van de leeromgeving
Hieronder staat een lijst met eigenschappen van het huidige opleidingssysteem. Zou je voor ons deze eigenschappen op volgorde van belangrijkheid kunnen zetten, gegeven de doelen die je net beoordeeld hebt? Let wel, de prioriteit die je aan een eigenschap toekent houdt direct verband met de doelen die je al dan niet wilt halen!
7. Achter de eigenschap die je het belangrijkst vindt zet je een 1, achter de minst belangrijke eigenschap 16. Gebruik alle getallen van 1 t/m 16 om de volgorde van belangrijkheid aan te geven. Doe het snel, je eerste intuïtie is meestal het beste.
De cursusstof wordt door middel van een kennis assessment op de cursist afgestemd. 1
Cursisten kunnen niet overleggen zodat zij ongestoord kunnen leren. 5
Cursisten kunnen overleggen zodat zij elkaar kunnen ondersteunen tijdens het leren. 4
Vragen over de cursus(stof) worden openbaar behandeld en kunnen worden beantwoord door cursisten en docenten. 2
Cursisten kunnen 24x7 vragen stellen aan een docent maar krijgen niet direct antwoord. 9
Cursisten kunnen zelf de tijd indelen waarop zij cursussen volgen. 16
Cursisten hebben beperkte tijd toegang tot het cursusmateriaal zodat zij worden gestimuleerd om de cursus op tijd af te ronden. 13
Kennis en ervaringen van cursisten worden vastgelegd voor hergebruik. 12
Cursussen zijn van langere duur en behandelen een compleet onderwerp. 3
Cursisten kunnen op beperkte vaststaande momenten vragen stellen aan een docent en krijgen direct antwoord. 6
Cursussen zijn van korte duur en behandelen een deelonderwerp. 7
Cursisten moeten een door ConQuaestor opgesteld cursusrooster volgen. 8
Kennis en ervaringen van cursisten worden niet vastgelegd en blijven ieders exclusieve eigendom. 10
Vragen over de cursus(stof) worden privé behandeld en kunnen alleen worden beantwoord door docenten. 11
Cursisten hebben onbeperkte tijd toegang tot het cursusmateriaal zodat dit na de cursus kan worden gebruikt als referentie. 14
Cursisten kunnen zelf een selectie maken uit een breed aanbod van cursusstof. 15
annuleren
klaar
Internet

4 Valence and Agreement

Abstract

This chapter attempts to contribute to a general theory of requirements change from a goal-oriented and viewpoints-driven angle.[1] To practitioners, this knowledge is relevant to anticipate changes in certain types of requirements, which may shorten the project's timeline, reduce costs, and increase product quality. Initially, I followed the common assumptions that what should be on a system is demanded by goals to achieve and what should not be on a system is demanded by goal states to avoid. However, requirements engineering of a diversity of systems (capacity and warehouse management, COTS PCs, and a Braille mouse) revealed that must requirements are predicted by goals to avoid (!) and won't requirements by goals to approach (!). Expectations about the positive or negative impact (valence) of requirements on goals played a moderating role. I unfold the gradual discovery of this "goals-to-requirements chiasm" (CHI-effect or χ-effect), claiming that variability in agreement to positive or negative requirements is predicted by goals of opposite polarity. Whether the χ-effect occurred or not, depended on the alignment of stakeholder viewpoints on goals and requirements. Comments from practitioners are included.

Keywords: Requirements, Validation, Change, Viewpoints, Goals, Measurement

4.1 Introduction

While a number of new printers were put into service at a New York business company, the management insisted on installing anti virus software although those printers never would communicate with the outside world. Jo Geraedts, head of the Industrial Design Department at Océ-Technologies, told me this story after being confronted with the results of my studies (personal communication, Nov. 11, 2004). This story exemplifies how a must requirement (i.e. protection) was predicated by a situation that stakeholders feared most (here, virus infection). It is precisely this relationship – goals to avoid direct the changes in what must be on a system – which underlies one of the hypotheses that I defend in the present chapter.

[1] This chapter is based on Hoorn et al., (2005) and Hoorn et al., (2006).

A statement of Arco van Nieuwland illustrates another hypothesis that I maintain. At that time chief executive at Exact Software, Van Nieuwland said (personal communication, Nov. 17, 2004): “In Europe, we won’t migrate e-Synergy [the software platform of Exact] to Linux because we want to keep that market, which is dedicated to Microsoft.” The desired goal was to preserve the European market for e-Synergy, which designated Linux as a won’t requirement - in spite of its technical advantages.

Yet, I started my investigations from the common assumption that agreement to requirements that are a must can be predicted from goals to achieve with the system. In this line, I also expected that agreement to won’t requirements could be predicted by goal states that stakeholders wanted to avoid. The idea was to know the goals and concerns of the system’s stakeholders in advance. When those goals were affected by, for example, a market event, I wanted to know how far agreement to the related requirements would change. Research question 1, therefore, was whether agreement to must requirements could be predicted from goals stakeholders wanted to approach while controlling for the effects of won’t requirements and goals to avoid. Research question 2 was whether goals stakeholders wanted to avoid predicted agreement to won’t requirements, while controlling for the influence of must requirements and goals to approach.

However, things turned out another way. I tested four cases that rendered five data sets and only in one out of five could I establish the relationship of goals to achieve with must requirements in coalition with goals to avoid with won’t requirements. In the four other data sets, I found the reverse relationship, which led me to articulate the *goals-to-requirements chiasm* (CHI-effect) or χ-effect for short. In brief, the χ-effect predicts that requirements most susceptible to change during a software-development track are governed by stakeholders’ goals that are inversely related. That is, goals to avoid with a system regulate changes in agreement to must requirements, whereas goals to achieve regulate changes in agreement to won’t requirements.

This finding is important to practitioners because it may provide a different focus to the requirements analysis. First, the ‘negative side of things’ seems to count as well. To understand a change request, won’t haves and goals to avoid are as important as must haves and goals to achieve. Second, when changes in the must requirements occur, one could look at the goals to avoid for the reasons why. Vice versa, when a won’t requirement changes, the reasons should probably be sought in goals to achieve with a system. Third, due to the systematic and statistical approach in this paper, results have a certain degree of reliability (they could be repeated, even in small groups) and add to a general theory of requirements change. In practice, statistically less intensive approaches could be applied to explore the proposed relationships in specific software-development cases.

In the remainder, I provide the theoretical background for the type of goals I studied and why I wanted to look at ‘negative’ requirements as well. I also

introduce the concept of valence. The hypotheses set up at the end of Section 4.2 are then tested with five data sets from four real life cases at the police, provincial government, a knowledge institution, and two blind schools. I used structured questionnaires and analyzed the results with multivariate analyses of variance and multivariate linear regression. In the Discussion (Section 4.7), I argue how the results contribute to the overall understanding of requirements and what the implications may be for practitioners.

4.2 Theory

4.2.1 The Type of Goals

In the area of goal-oriented requirements engineering (RE) (e.g., Van Lamsweerde, 2004), the cause of requirements change, requirements evolution (Alves & Finkelstein, 2002), or requirements development (Robinson & Pawlowski, 1999) is sought in the goals that stakeholders want to achieve with the system or the concerns they may have with it. "Goals are ... essential elements for managing requirements evolution" (Letier & Van Lamsweerde, 2001). Goals can range from high-level strategic mission statements to low-level operational targets that should be achieved with the system (Letier, & Van Lamsweerde, 2001). Goals are supposed to be more stable than the requirements that help reaching them (Van Lamsweerde & Letier, 2000). Moreover, the higher-level a goal is (e.g., a strategic business goal), the more stable the respective requirements will be (Anton, Cracken, & Potts, 1994; Alves & Finkelstein, 2002). Thus, the reasons for requirements change should be sought in a change of lower-level goals, such as improving a work process (e.g., higher efficiency, less costs), or advancing system performance, security, and reliability.

4.2.2 Valence

When stakeholders are involved in developing a system, they are – whether intentionally or not – also busy designing the future situation of their business or work environment (Chapter 2, BTM2). Therefore, they make evaluations of how much a requirement, once implemented as a feature of the system, will impact their goals (cf. Lehman, 1996).

In goal-driven RE, system development is centered on the stakeholders' concerns (Anton et al., 2001; Alves & Finkelstein, 2003). Concerns are basic to emotions, which in turn motivate people to undertake action (Frijda, 1986). In this line, I assume that the requirements on the new system are judged for their usefulness or relevance to potentially satisfy or harm the stakeholder's concerns, goals, or motives. Positive expectations about the future situation result from requirements that promise a match, the actual or expected satisfaction of concerns. Negative expectations result from requirements that promise a mismatch, the actual or expected obstruction of realization of goals and concerns (Frijda, 1986, p. 277). Frijda (p. 207) points out that valence refers to the implied outcome of the event: The intrinsic attractiveness or repulsiveness. In

other words, valence (also Sutcliffe, 2002) refers to the expected match or mismatch between the potential gratification for or obstruction of stakeholder concerns and the possibilities or impossibilities offered by the new situation.

Stakeholders expect positive or negative consequences of the system for achieving their goals (cf. Technology Acceptance Model, Davis, 1989). Whether stakeholders expect that a proposed feature will support or obstruct their goals may have an impact on the level of agreement or disagreement to a requirement. When the business environment changes, the direction of valence towards the future system may change accordingly, thus triggering a change of requirements.

I suspected that valence could interfere with the agreement to requirements. A requirement may technically satisfy a goal. However, if people nevertheless expect negative effects on their work situation, the agreement to that requirement may be low. I assumed that the assessment of valence would be a necessary step to make a judgment about a requirement. In other words, I believed that valence would be a mediator of agreement in between goals on the one hand and requirements on the other (Chapter 2, Figure 7).

4.2.3 Not Only Must Haves

Although practitioners often work from a MuSCoW list,[2] the won't requirements are often put aside as irrelevant for further analysis. The focus is on the must haves, understandably, to help achieve the stakeholders' goals. However, whereas goals specify desired situations, so-called "obstacles" designate goal states that are undesirable but yet possible (Potts, 1995; Van Lamsweerde & Letier, 2000). Apart from achieving goals, there is also an "avoid-mode" (Robinson & Pawlowski, 1999). Thus, must haves may be important to achieve goals stakeholders want to approach, yet, won't haves are important to construe what stakeholders want to avoid with the system (e.g., instability, complexity, or a Linux shell for the e-Synergy platform). When a business model changes, the won't requirements may change just as well as the must requirements.

4.2.4 Variability in Agreement

When business goals change and the requirements change accordingly, the once agreed-upon requirements are often disagreed-upon in the new situation. If we know which goals have changed it should be possible to predict the level of agreement to the related requirements from the level of agreement to the (changed) goals. I suspected that requirements that raise the most conflicts among stakeholders are also most vulnerable to change. Such requirements should show more variability in the level of agreement (from agree to disagree) than requirements that raise no conflicts (a ceiling effect of either agree or disagree). Thus, I wished to investigate which type of goals (those to ap-

[2] Requirements that Must be, Should be, Could be, or Won't be on the system (eRA, 2002).

proach or those to avoid) best predicted the variability in the level of agreement to must or won't requirements. My best guess was that

- (H1) goals to approach predict agreement to the must requirements through the mediation of positive outcome expectancies (valence support)

In opposition, I assumed that

- (H2) goals to avoid predict (dis)agreement to won't requirements, mediated through negative valence (valence obstruct)

I tested these hypotheses in four different business cases that rendered five sets of data (one split file).

The remainder of this chapter is organized as follows. Study 1 in Section 4.3 evaluates H1 and H2 from two different viewpoints. In the personal view, the straightforward approach-to-must and avoid-to-won't relationships were established. However, from a business viewpoint a different constellation occurred that countered H1 and H2. The studies in Sections 4.4 to 4.6 attempted to repeat the results of the personal view in Study 1 but I systematically encountered another constellation (the χ-effect), in line with the business view. Therefore, I redefined the hypotheses into a set of precise test predictions (Section 4.5) that should be confirmed to speak of the χ-effect. The Discussion in Section 4.7 offers post hoc explanations of why initially I found the straight relationships and how those findings can be brought in line with the results of the subsequent studies. Section 4.7 closes with comments from and recommendations to practitioners in RE.

4.3 Study 1: Viewpoints on Requirements and Goals Unaligned

Together with Evelien Kok, I investigated the design of a Capacity Management System (CMS) at the Dutch police force (Chapter 3, Section 3.2). The corps management set up the requirements on this system, meant for scheduling police tasks and allocating personnel. We had 33 novice users, young officers, judge these management-viewpoint requirements (Sommerville & Sawyer, 1997) from two perspectives: A personal and a business point of view.

Our first research aim at that time was to test in how far the future end-users agreed to the management-viewpoint requirements from a personal point of view. The second research aim was to investigate the agreement to the management-viewpoint requirements while the end-users adopted a business point of view. The results forced us to split the data file into a set for the personal and a set for the business view to test H1 and H2.

4.3.1 Method

4.3.1.1 Participants and Experimental Design

See Chapter 3, Section 3.2.1.1, for a description of the relevant stakeholders. The sample of officers we took was split into two groups. One group of offi-

cers worked from the perspective that the management-viewpoint requirements should satisfy Personal Goals (n= 16). The other group worked on the same list of management-viewpoint requirements from the perspective that Business Goals should be satisfied (n= 17). This was the between-subjects factor of Stakeholders' View.

We also devised a within-subjects factor of Stakeholders' Needs. This was a nested factor of Requirements (Must vs. Won't) vs. Valence (Support vs. Obstruct) vs. Goals (Approach vs. Avoid). Goals to Approach or Avoid varied within Stakeholders' View and thus could be Personal or Business.

4.3.1.2 System

See Chapter 3, Section 3.3.1.2.

4.3.1.3 Procedure

See Chapter 3, Section 3.3.1.3.

Sample requirements in the CMS case

- Must (Business)
  - Registering work activities in advance
  - Digital registration of results of activities
  - Detailed work plans
  - Planning one year ahead
  - Planning based on skills of the personnel
  - Spending least possible amount of time on inputting data
  - Fixed holidays, fixed schedules
  - Personal portfolio of courses and training taken
- Won't (Business)
  - Work activities reported through face-to-face briefing
  - Results in written day reports
  - Global work plans
  - Planning one month ahead
  - Planning based on available personnel
  - Wasting time on data input
  - Schedules can be continuously changed
  - Flexible holiday planning
  - Courses and training remain unregistered

Sample goals in the CMS case

- Approach (Personal)
  - Intensive working relationship with my chief

    - Being my chief of service
    - Being involved with colleagues
    - Have freedom to show initiative
    - Insight in what colleagues do
    - Spend most of my time on the streets (help civilians)
    - Time to help my colleagues
    - Maintain my scheduling privileges

- Avoid (Personal)
    - Less contact with the chief
    - Frustrate the chief
    - Being indifferent about the colleagues
    - Only do what was planned
    - What colleagues do goes unnoticed
    - Spend most time behind the desk (form filling)
    - Everybody solves their own problems
    - All scheduling privileges are dropped

- Approach (Business)
    - Minimize absences
    - Make professional impression
    - Keeping the performance contract with the government
    - Save money on allocating personnel
    - Peace on the work floor
    - Keeping Internal Affairs out
    - Planning schedules on time
    - Cooperation between precincts

- Avoid (Business)
    - Stabilize absences
    - Make amateurish impression
    - Breaking the performance contract
    - Spend more money on personnel
    - Irritation on the work floor
    - Interference by Internal Affairs
    - Planning schedules late
    - Each precinct works on its own

4.3.1.4 Measurements

Scale construction

For the uninformed reader, I want to introduce the notions of structured questionnaire design (Dillman, 1999), scales, indicative and contra-indicative

items, and faceted scales (Guttman, 1954; 1965). In the next section I explain how the measurements were done in practice.

Scales measure a concept or construct that is not immediately visible in the concrete world (e.g., stakeholder goals). Scales consist of multiple items that more-or-less cover a variety of aspects of 'stakeholder goals' (e.g., efficiency, cost-effectiveness, or fun). Together, the items cover the abstract concept of stakeholder goals not only from the positive side ("E-mail is fast") but also from the negative side ("E-mail is slow"). Such statements form the indicative and contra-indicative items on the scale, respectively. Each item is scored for agreement on rating scales. Taken together, the various items on a scale control different interpretations of what 'stakeholder goals' might mean. Faceted scales (Guttman, 1954; 1965) systematically combine more single (sub) scales (e.g., requirements plus valence plus goals). Thus, a statement from a faceted scale can be formulated as a requirements statement (e.g., "Automated input helps me to do my work properly"). "Automatic input" is the must requirement, "helps me" induces positive valence, and "work properly" is a goal to approach. Each item is part of a larger set of statements that systematically combine, for example, the positive and negative aspects of the respective sub scales to see their different impact on agreement.

Together, items on the faceted scale Stakeholders' Needs combined a requirement with a certain valence to a goal. Items on the scale Stakeholders' Needs followed the structure:

<Requirement (must or won't have)> has <Valence (supports or obstructs)> towards a <Goal (that you want to approach or want to avoid) >

By systematically combining the three sub scales, I produced eight categories of items. For each category, 3 variants were prepared, resulting in 24 items on the scale Stakeholders' Needs.

1. Must requirement – supports – goal to approach (× 3)
2. Must requirement – supports – goal to avoid (× 3)
3. Must requirement – obstructs – goal to approach (× 3)
4. Must requirement – obstructs – goal to avoid (× 3)
5. Won't requirement – supports – goal to approach (× 3)
6. Won't requirement – supports – goal to avoid (× 3)
7. Won't requirement – obstructs – goal to approach (× 3)
8. Won't requirement – obstructs – goal to avoid (× 3)

Each sub scale, then, (Requirements Must, Requirements Won't, Valence Support, Valence Obstruct, Goals Approach, Goals Avoid) had 12 items (that is, 3 items coming from 4 item categories). Scale analysis (e.g., Section 4.3.2.1) was performed on the 12 items per sub scale.

Take notice that all items were presented to the stakeholders in the various studies as affirmative statements to avoid answering biases and response confusion from using linguistic negations (Dillman, 1999). That is, the won't requirements as well as the goals to avoid were put as desirable things and the stakeholders were expected to disagree to these won't (put as must) requirements and the avoid (put as approach) goals.

For all the studies reported in this paper, the above structure and rationale was followed to create questionnaire items on the various Stakeholders' Needs scales. For different purposes, each Requirements Engineering questionnaire (RE*quest*) featured more and other scales than the Stakeholders' Needs scale alone. However, these extra scales have less relevance to the hypotheses tested here and will only be mentioned if necessary.

Scale construction in the CMS case

Stakeholders' Needs consisted of requirements statements that systematically connected goals with the management-viewpoint requirements, while putting positive or negative valence to them. Together, the Requirements, Valence, and Goals formed a so-called faceted scale. Note that in one group of officers, the goals were personally oriented and in the other business oriented. For the questionnaires and additional information, see Chapter 3 (Section 3.2.1.4 and Appendix 3.1).

The items on the Stakeholders' Needs scale followed the structure set out in the previous section. In this way, eight different categories of items (i.e. requirements statements) were established. Examples are (translated from Dutch):

1. That schedules are definite 48 hours in advance helps a relaxed work pace
2. That schedules are definite 48 hours in advance frustrates a relaxed work pace
3. That schedules are definite 48 hours in advance decreases the time pressure
4. That schedules are definite 48 hours in advance increases the time pressure
5. That schedules can change continuously helps a relaxed work pace
6. That schedules can change continuously frustrates a relaxed work pace
7. That schedules can change continuously decreases the time pressure
8. That schedules can change continuously increases the time pressure

The items on the Stakeholders' Needs scale were scored for agreement using 6-point rating scales (0= completely disagree, 5= completely agree).

For each type of item we made three exemplars, resulting in 24 items on the Stakeholders' Needs scale. From this structure, 6 unipolar sub scales could be extracted for the analysis: Requirements Must, Requirements Won't, Valence Support, Valence Obstruct, Goals Approach (either Personal or Business), and Goals Avoid (either Personal or Business).

The items were tested by focus groups for readability, wording, and whether their contents made sense to people working in the field. After the necessary repair work, items were again inspected by a focus group, after which we considered them ready for the main test.

4.3.2 Analysis and Results

Table 1: Motivation of most important statistical techniques used in the case studies.

Statistical technique	**Explores**	**To answer the question**
Standardized Cronbachs's alpha	Average correlation of items within a scale (internal consistency reliability)	Can I trust what I measured?
Corrected Item-Total Correlations	Correlations of a single item with the sum of all other items	Can I trust what I measured? Are items unambiguous or do they belong to more scales? Thus, items should not correlate strongly with other scales
Pearson correlations (pmcc)	The degree to which two variables are related	
(Multivariate) Analysis of Variance - (M)ANOVA	The statistical significance of the differences among the mean scores of two or more groups on one or more variables	Does the level of agreement to must requirements really differs from the won't requirements? If not, requirements intended as won't are perhaps must and v.v. Likewise for goals to approach vs. to avoid and for valence support vs. obstruct. Are there differences in agreement within and between a business or a personal view? Do background variables such as sex and age affect the level of agreement?
Multiple linear regression analysis	The conditional expected value of one variable (the dependent) given the combined effect of multiple other variables (the predictors). Incrementing the predictor(s) supposedly leads to a fixed amount of increment of the dependent (no curvilinear relations)	Are the hypotheses confirmed or rejected? To what extent can I predict agreement to requirements from agreement to goals? What is the modality (role) of valence in predicting agreement to requirements? How much does valence contribute to the level of agreement to requirements?

After the completed questionnaires were returned, the data were entered in an SPSS 11.0 data matrix for statistical analysis. In Section 4.3.2.1, I evaluate the sub scales of the Stakeholders' Needs scale for psychometric quality. In

Sections 4.3.2.2 and 5.3.2.3, I check how much the officers agreed to the goals and management-viewpoint requirements from their personal perspective. To test H1 and H2, I explored the relations among the different sub scales with multiple linear regression analyses in Section 4.3.2.4. In-depth details about the statistical procedures followed and intermediate results can be found in (Hoorn & Kok, 2005, Tech. Rep. [CD]). Table 1 shows an overview of statistical techniques employed in the current paper and the function of each technique.

4.3.2.1 Scale Analysis

I assessed the psychometric quality of the 6 unipolar sub scales Requirements Must, Requirements Won't, Valence Support, Valence Obstruct, Goals Approach, and Goals Avoid (cf. Table 1, first row). Each sub scale had 12 items each (see Section 4.3.1.4). I tested whether items correlated sufficiently with their own scale by means of Corrected Item-Total Correlations and Standardized Cronbach's alpha (indicating reliability). The degree to which items did not correlate with other scales was tested with Pearson correlations.

Item selection was a trade-off among several criteria. I wanted to establish as many items on a sub scale as possible with a minimum of 2, provided that Standardized Cronbach's alpha for a scale was at least > .60, preferably > .70, and that items showed the lowest correlations possible with other scales. SD of items should be around 1. I also wanted the skewness of items and scale < .70 and if present, I removed so-called leverage points.[3]

I correlated each item with its own sub scale (with the item removed) and with the other sub scales. In many cases, items were more highly correlated with another sub scale than with their own sub scale. Probably, this is because the items on the Stakeholders' Needs scale explicitly related requirements, valencies, and goals, which may explain the relatively strong interdependency of sub scales. Based on these results and additional item analyses, the psychometrically weak items were eliminated from their sub scales. Each item on the shortened scales was again correlated with its own sub scale (with the item removed) and with the other sub scales.

The thus revised sub scales in the Personal View had a length of 2 to 4 items and showed a Standardized Cronbach's alpha between .77 and .90, which is fine. The revised sub scales in the Business View had a length of 2 to 4 items. Standardized Cronbach's alpha's were between .62 and .91, which is acceptable to good.

4.3.2.2 Agreement to CMS Requirements and Goals

Before exploring the relations among requirements, valence, and goals, I checked to what degree the officers actually agreed to the management-viewpoint requirements and to the personal and business goals we gathered

[3] Leverage points are values extremely distant from the center of the sampled predictor values.

(cf. Table 1, second row). Moreover, I wanted to see whether socio-demographic variables such as age, sex, and function influenced the results.

To do so, I treated the Stakeholders' Needs scale as a nested factorial design (within-subjects) of the 3-leveled factor Scales (Requirements vs. Valence vs. Goals) and the 2-leveled factor Item Type (Indicative vs. Contra-indicative). In view of this setting, 6 within-subjects (dependent) variables were calculated from the 2 or 3 items per sub scale: The grand mean level of agreement to Requirements (Must vs. Won't) vs. Valence (Support vs. Obstruct) vs. Goals (Approach vs. Avoid).[4] This was done for the Personal as well as the Business View. The results are in Figure 1.

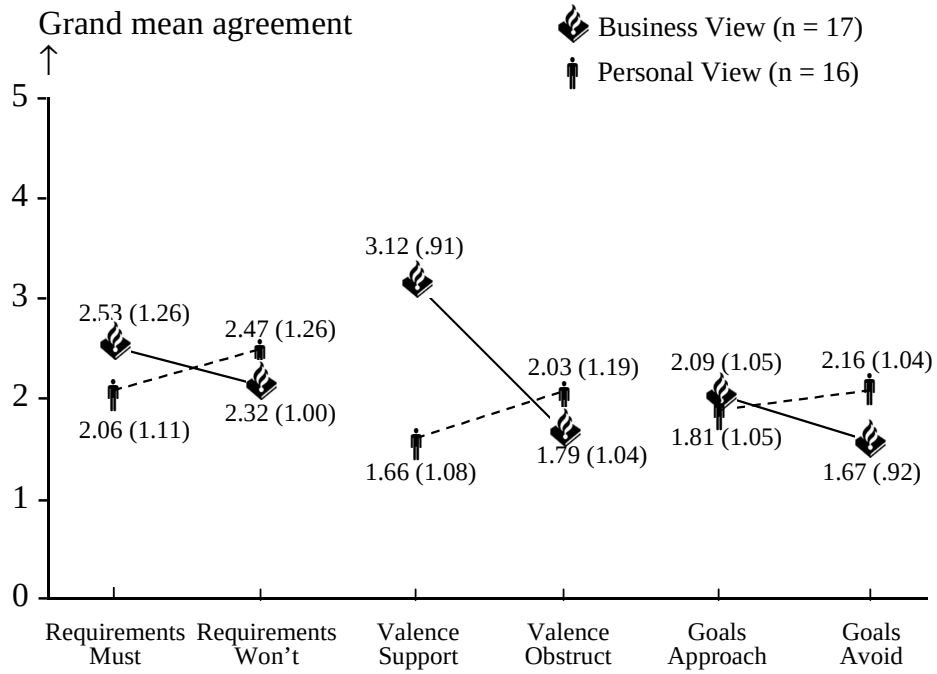


Figure 1: Grand mean averages of agreement to Stakeholders' Needs (Must, Won't, Support, Obstruct, Approach, Avoid) from a Business and a Personal View. Standard deviations are in parentheses (N= 33).

I then ran a 2*2*3 MANOVA of Stakeholders' View (Personal vs. Business) (between-subjects) by Item Type (Indicative vs. Contra-indicative) (within-subjects) and Scales (Requirements vs. Valence vs. Goals) (within-subjects) on the grand mean average level of agreement.[5]

Two significant interactions were established. The interaction depicted in Figure 2 between Stakeholders' View (Business vs. Personal) and Item Type

[4] Grand means are averages across the individual means.

[5] Note that the GLM > Repeated measures option in the new releases of SPSS is more-or-less similar to the MANOVA procedures available through the syntax editor. The latter option was used in all the studies reported here.

(Indicative vs. Contra-indicative) ($F_{(1,30)}$= 13.76, p= .001, η_p^2= .31,[6] parameter coefficient= 1.58, t= 3.71, p< .001) showed that in the case of the Business View, the level of agreement to Indicative items (Must, Support, and Approach) (M_{MSAp}= 2.58, SD= 1.07) was always higher than to Contra-indicative items (M_{WOAv}= 1.92, SD= .98). Contrariwise, in the Personal View, the level of agreement to Contra-indicative items (Won't, Obstruct, and Avoid) (M_{WOAv}= 2.22, SD= 1.24) was always higher than to Indicative items (M_{MSAp}= 1.84, SD= 1.08).

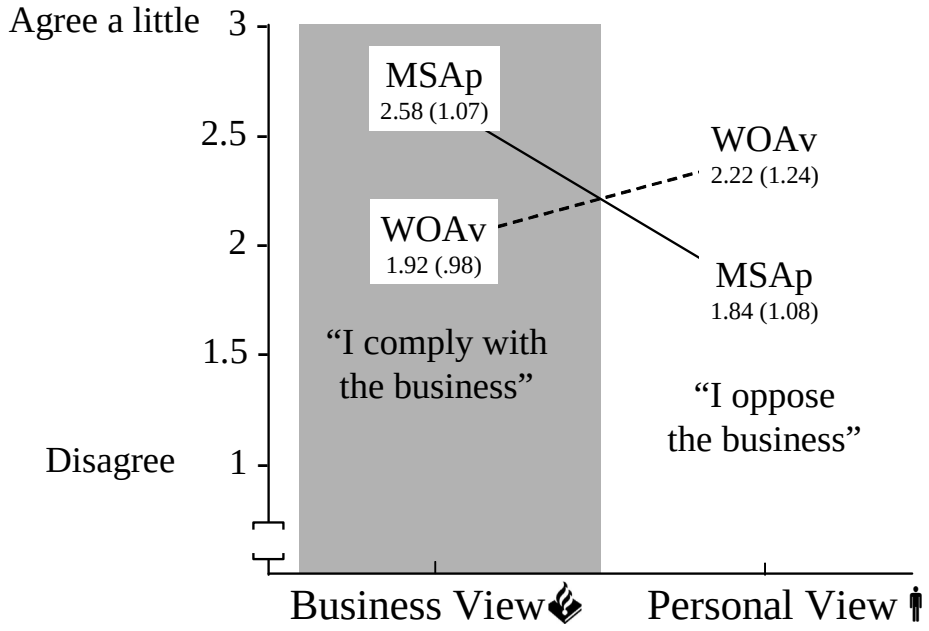


Figure 2: Grand mean level of agreement to Must, Should, and Approach (MSAp) items versus Won't, Obstruct, and Avoid (WOAv) items for two different views: Business and Personal. Standard deviations are in parentheses (N= 33).

Another significant interaction (depicted in Figure 3) occurred between Stakeholders' View (Business vs. Personal) and Scales (Requirements vs. Valence vs. Goals). It pointed toward an advantage of Valence over Goals in increasing the level of agreement in the Business View but the reverse occurred in the Personal View (Pillai's Trace= .35, $F_{(2,29)}$= 7.92, p= .002). The parameter Stakeholders' View * (Valence vs. Goals) showed that at an α-level according to Bonferroni ($\alpha = .05/3 \approx .017$), Valence in the Business View ($M_{Valence}$= 2.46, SD= .97) evoked higher levels of agreement than Goals in the Business View (M_{Goals}= 1.87, SD= .98). Reversely, Valence from a Personal point of View ($M_{Valence}$= 1.84, SD= 1.14) evoked lower levels of agreement than Goals in the

[6] Partial eta squared (η_p^2) indicates the effect size or the proportion of variance explained by a (combination of) variable(s).

Personal View (M_{Goals}= 1.98, SD= 1.15) (parameter coefficient= .73, t= 4.05, p< .001, η_p^2= .35) (Figure 3).

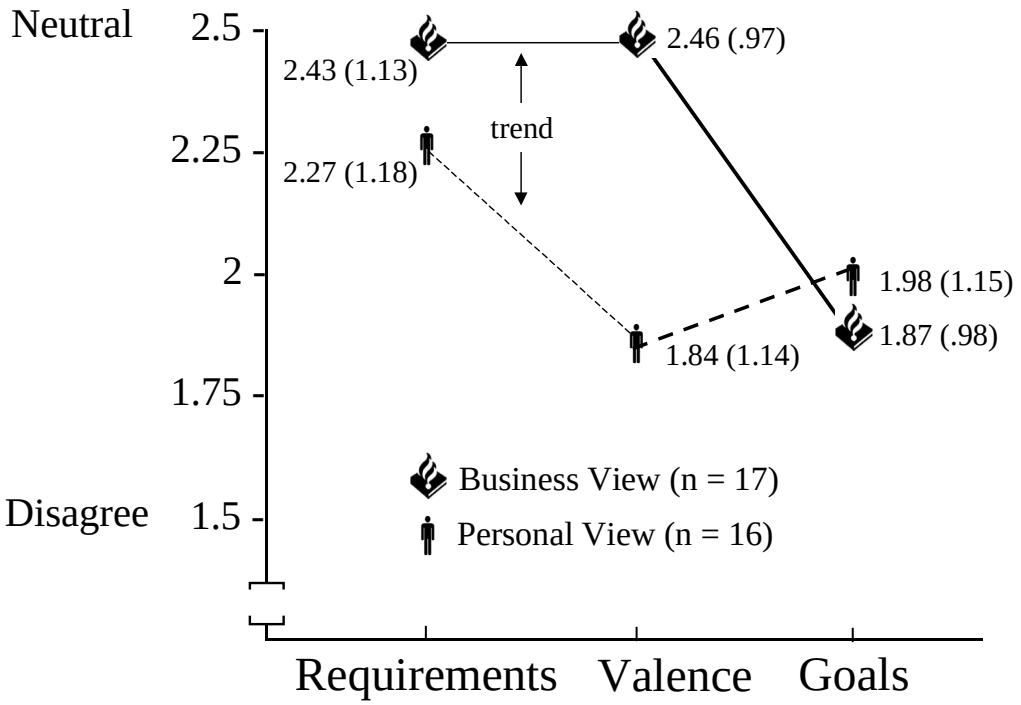


Figure 3: Grand mean level of agreement to Requirements, Valence, and Goals from a Business and a Personal View. Bold lines indicate significance at $\alpha \approx .017$, thin lines at α = .05. Standard deviations are in parentheses (N= 33).

If the lenient rejection area of α= .05 is accepted, the parameter Stakeholders' View * (Requirements vs. Valence) underscored the trend that Requirements in the Business View (M_{Reqs}= 2.43, SD= 1.13) elicited lower levels of agreement than Valence in the Business View ($M_{Valence}$= 2.46, SD= .97). In opposition, Requirements from a Personal point of View (M_{Reqs}= 2.27, SD= 1.18) yielded higher levels of agreement than Valence in the Personal View ($M_{Valence}$= 1.84, SD= 1.14) (parameter coefficient= -.45, t= -2.16, p= .039, η_p^2= .13).

I repeated the entire analysis with Sex (2), Function (2), Years in Service (4), Years in present Function (3), and Number of Years at the Academy (3) as fixed factors and controlled for Age as covariate. However, none of the effects of these background variables were significant (Hoorn & Kok, 2005, Tech. Rep. [CD]).

4.3.2.3 Discussion of the Effects on Agreement in the CMS Case

The significant interaction between Stakeholders' View (Business vs. Personal) and Item Type (Indicative vs. Contra-indicative) showed that from a

Business point of View, the level of agreement to Must, Support, and Approach (MSAp) items was higher than to Won't, Obstruct, and Avoid (WOAv) items. From a Personal point of View, however, the level of agreement to Must, Support, and Approach (MSAp) items was lower than to Won't, Obstruct, and Avoid (WOAv) items (Figure 2).

The left panel of Figure 2 suggests that when these police officers took the point of view of the business, they complied with the business. They more-or-less agreed to what the corps managers proposed to implement on the CMS (e.g., precise administration of absences) and to the business goals to achieve with that (e.g., reduce costs on personnel). There were also features the corps managers did not want to have on the future CMS (e.g., schedules that can continuously change) and goal states they did not want to get into (e.g., irritation on the work floor). When in the questionnaire these requirements and goals were presented as things the future CMS would have and as desirable goals, these police officers disagreed, as would be expected from a business point of view. Moreover, the police officers had positive expectations when the managers had positive expectations and negative expectations when the managers did.

However, the attitude changed drastically when requirements on the CMS and goals were judged from a personal perspective (Figure 2, right panel). Personally, the police officers wholeheartedly disagreed with their managers. The won't requirements of the managers raised more agreement with the officers than the managers' must requirements. The same happened for the goals to approach and those to avoid. In other words, the personal goals that were proposed to these officers as things to achieve (e.g., "to be of service to my chief") were what they wanted to avoid. What was presented as personal goals to avoid (e.g., "less contact with the chief"), these officers wanted to achieve. Obviously, what the managers thought that should have been on the CMS to achieve these personal goals, was rejected by the officers as won't have requirements. Likewise, what the managers proposed as won't requirements, these officers saw as must haves. Additionally, the officers disagreed to what the managers thought would have a positive impact on the work environment and the officers agreed to what would obstruct them in working with the CMS.

In other words, if a systems engineer asked these future users of the CMS personally, it would turn out that what must be for the managers with regard to business goals, won't be for the officers with regard to their personal goals. What these managers won't have with regard to business goals is what the officers deem a must for their personal goals: This position designates a true requirements-analysis rift (Chapter 3). However, this rift would not come to the fore if the engineer did not ask these officers personally, because then the officers would lawfully join the position of their corps.

The significant interaction between Stakeholders' View (Business vs. Personal) and Scales (Requirements vs. Valence vs. Goals) showed that Valence (whether positive or negative) in the Business View provoked more agreement

than Valence from a Personal View (Figure 3). For goals, it was the other way round. Personal Goals (whether to Approach or Avoid) raised more agreement than Business Goals. As a trend, the Requirements (whether Must or Won't) proposed by the corps management evoked more agreement from a Business ("It's their system") than from a Personal View ("What should I say about it?").

The significant interaction between Stakeholders' View (Business vs. Personal) and Scales (Requirements vs. Valence vs. Goals) displayed in Figure 3 illustrates that these police officers disagreed more to the Business Goals of their managers (M_{Goals}= 1.87) than to the Personal Goals we obtained from our ethnography (M_{Goals}= 1.98). This was irrespective of whether these Business and Personal Goals were desirable or not. However, from what these officers expected of the CMS for their future work situation (Valence), they foresaw little impact on their Personal situation ($M_{Valence}$= 1.84), for better or for worse. Yet, the future CMS was expected to impact strongly the future Business situation with regard to planning and control ($M_{Valence}$= 2.46), whether this impact was positive or negative. Finally, in agreeing to the requirements (whether must or won't), these officers followed the propositions of the corps (Business View: M_{Reqs}= 2.43), which was valued higher than their Personal View on the requirements (M_{Reqs}= 2.27).

These results seem like a clear-cut victory for the corps management. The future users of the CMS foresaw little change in their personal work situation (personal valence low), the business was expected to undergo the necessary change (business valence high), and the users were more-or-less neutral to what should be on or off the future CMS (requirements neutral). However, the future users did not subscribe to the business goals that should be achieved or avoided with the future CMS and they valued their personal goals higher (Price & Cybulski, 2004). Taking this result together with the results exhibited in Figure 2, it seems that these officers did not see the connection between what they wanted and what the CMS was going to offer. They did not see much impact on their own work, but did agree that the business could benefit from the future CMS. Again, the requirements-analysis rift (Chapter 3) became visible: The requirements were considered something of the business whereas goals were something of the person but the persons did not realize that their future use of the system would probably facilitate the business but frustrate their personal goals at work.

I observed almost diametrically opposed levels of agreement between the Business and Personal View on the same list of management-viewpoint requirements and the related goals. I therefore decided to split up the data set to verify the hypotheses H1 and H2 separately for the Personal and the Business View. Section 4.3.2.4 scrutinizes the relations between goals and requirements for the Personal View. Section 4.3.2.6 does the same for the Business View.

4.3.2.4 CMS: Management-viewpoint Requirements in Relation to Personal Goals

Originally, I advocated the common view that (H1) Goals Approach predicts agreement to Requirements Must. I added the idea that Valence Support would serve as a mediator. In addition, I assumed (H2) that Goals Avoid predicts agreement to Requirements Won't and that Valence Obstruct would serve as the mediator. To test these hypotheses, I performed several multiple linear regression analyses (method Enter) on the grand mean average agreement to the sub scales of Stakeholders' Needs (cf. Table 1, third row).

Explaining Requirements Must

To verify H1, a multiple linear regression analysis (method Enter) was performed on the Personal View data set. Requirements Must served as the dependent variable with three ordered sets of predictors. Requirements Won't was entered in the first step, Valence Support and Valence Obstruct in the second step, and Personal Goals Approach and Personal Goals Avoid in the third (Hoorn & Kok, 2005, Tech. Rep. [CD]).

The third set of predictors, Personal Goals Approach and Personal Goals Avoid, accounted for a significant amount of the Requirements Must variability, R^2= .75, R^2_{adj}= .63, $F_{(5,10)}$= 6.03, p= .008. On the basis of correlation-regression analyses, the relative importance of Personal Goals Approach and Personal Goals Avoid in predicting Requirements Must was assessed. It seemed that Personal Goals Approach was most strongly related to Requirements Must, standardized β= .86, t= 3.49, p= .006. Supporting this conclusion is the height of the standardized Beta coefficient and the strength of the correlation between Personal Goals Approach and Requirements Must partialling out the effects of all other predictors ($r_{partial}$= .74, r_{part}= .55). Personal Goals Avoid offered little or no additional predictive power beyond that contributed by the Personal Goals Approach measure.

Explaining Requirements Won't

With regard to H2, Requirements Won't served as the dependent variable with three ordered sets of predictors. Requirements Must was entered in the first step, Valence Support and Valence Obstruct in the second step, and Personal Goals Approach and Personal Goals Avoid in the third (Hoorn & Kok, 2005, Tech. Rep. [CD]).

The second set of predictors, Valence Support and Valence Obstruct, accounted for a significant amount of the Requirements Won't variability, R^2= .49, R^2_{adj}= .36, $F_{(3,12)}$= 3.85, p= .038. The third set, Personal Goals Approach and Personal Goals Avoid, however, predicted significantly the percent of explained variance of Requirements Won't over and above the Valence measures, R^2_{adj}= .77, R^2_{change}= .36, $F_{(2,10)}$= 11.45, p= .003. On the basis of correlation-regression analyses, the relative importance of Personal Goals Approach and Personal Goals Avoid in predicting Requirements Won't was assessed. It seemed that Personal Goals Avoid was most strongly related to Requirements

Won't, standardized β= 1.07, t= 4.14, p= .002. Supporting this conclusion is the height of the standardized Beta coefficient and the strength of the correlation between Personal Goals Avoid and Requirements Won't partialling out the effects of all other predictors ($r_{partial}$= .80, r_{part}= .52). Personal Goals Approach offered little or no additional predictive power beyond that contributed by the Personal Goals Avoid measure.

In other words, Valence did contribute to the Requirements Won't variability but Personal Goals Avoid overruled Valence Support and Valence Obstruct in predictive power.

4.3.2.5 Discussion of the Relation between Management-viewpoint Requirements and Personal Goals

The results of the multiple regression analyses in the Personal View seem to support the common assumption (H1) that striving for positive goals directs positive requirements (Personal Goals Approach explained Requirements Must, β= .86, $r_{partial}$= .74, r_{part}= .55). Likewise (H2), avoiding negative situations determines what should not be on the system (Personal Goals Avoid explained Requirements Won't, β= 1.07, $r_{partial}$= .80, r_{part}= .52). In itself, this is a plausible result in that a computer system should have, for example, a keyboard for text entry or should not be so small that human hands cannot handle it.

Figure 4 illustrates this stance, showing that the status of valence was not what I expected. The idea was that valence assessment (whether requirements would work out for good or for ill) was necessary to reach a level of agreement. However, valence did not play such a mediating role because it did not flush out the predictive power of Personal Goals Avoid. Figure 4 draws the analogy between the role of valence as a moderator of agreement and the influence of the weather on someone's mood. A person can be happy because s/he got married (the cause) and that the sun is shining enhances but is not the cause of that mood. Similarly, a stakeholder may want to have an easy-to-handle system (goal to approach) and therefore wishes a Windows shell (must requirement). The prospect that Windows is error-prone, however, tempers her enthusiasm – valence obstruct is the moderator and not the reason why she yet agrees to Windows.

I could have left it here. In fact, the original hypotheses were confirmed with a slight alteration in the status of valence. However, the results of the Business View data set in combination with the requirements-analysis rift I found in the Sections 4.3.2.2 and 4.3.2.4 made me wonder if something else was at hand.

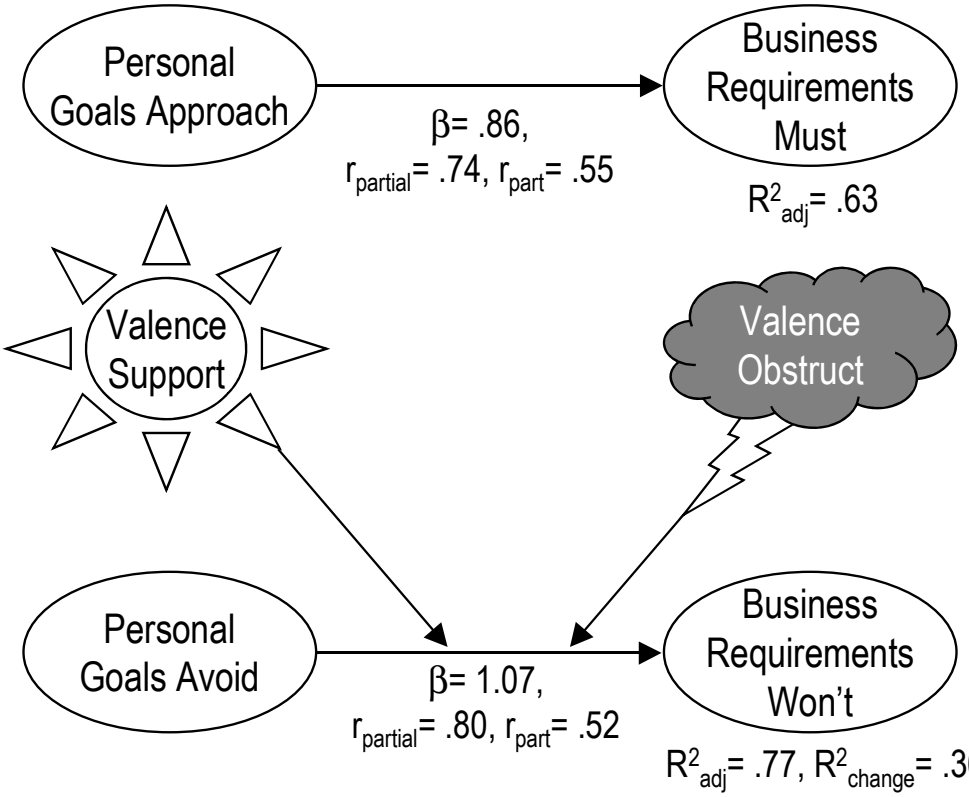


Figure 4: Straight relationships between Goals to Avoid and Won't Requirements as well as Goals to Approach and Must Requirements. Valence is not necessary to agree to requirements but moderates the relationship.[7]

4.3.2.6 CMS: Management-viewpoint Requirements in Relation to Business Goals

I followed the same approach as put forth in Section 4.3.2.4 except that I took the data set that was related to the Business Goals.

Explaining Requirements Must

Requirements Must was the dependent variable with three ordered sets of predictors. Requirements Won't was entered in the first step, Valence Support and Valence Obstruct in the second step, and Business Goals Approach and Business Goals Avoid in the third.

The second set of predictors, Valence Support and Valence Obstruct, accounted for a significant amount of the Requirements Must variability, R^2= .80, R^2_{adj}= .65, $F_{(3,13)}$= 7.93, p= .003. The third set, Business Goals Approach and Business Goals Avoid, did not significantly increment the percent of explained variance of Requirements Must, R^2_{change}= .09, $F_{(2,11)}$= 1.94, p= .189. I assessed the relative importance of Valence Support and Valence Obstruct in

[7] These are actually two sub models (Approach→Must and Avoid→Won't). Beta weights and correlations cannot be compared therefore.

predicting Requirements Must on the basis of correlation-regression analyses. Valence Obstruct was most strongly related to Requirements Must, standardized β= 1.03, t= 3.65, p= .003. This is underscored by the height of the standardized Beta coefficient and the strength of the correlation between Valence Obstruct and Requirements Must partialling out the effects of all other predictors ($r_{partial}$= .71, r_{part}= .60). Valence Support offered little or no additional predictive power beyond that contributed by the Valence Obstruct measure.

Explaining Requirements Won't

The dependent variable Requirements Won't had three ordered sets of predictors. Requirements Must was entered in the first step, Valence Support and Valence Obstruct in the second step, and Business Goals Approach and Business Goals Avoid in the third.

The second set of predictors, Valence Support and Valence Obstruct, accounted for a significant amount of the Requirements Won't variability, R^2= .77, R^2_{adj}= .71, $F_{(3,13)}$= 14.22, p= .000. The third set, Business Goals Approach and Business Goals Avoid, however, significantly incremented the percent of explained variance of Requirements Won't, R^2_{adj}= .89, R^2_{change}= .16, $F_{(2,11)}$= 11.88, p= .002. On the basis of correlation-regression analyses, the relative importance of Valence Support and Business Goals Approach in predicting Requirements Won't was assessed. It seemed that Business Goals Approach was most strongly related to the Requirements Won't, standardized β= .69, t= 4.59, p= .001. Supporting this conclusion is the height of the standardized Beta coefficient and the strength of the correlation between Business Goals Approach and Requirements Won't partialling out the effects of all other predictors ($r_{partial}$= .81, r_{part}= .38). Valence Support offered some additional predictive power (standardized β= .28, t= 2.60, p= .025, $r_{partial}$= .62, r_{part}= .21) but not beyond that contributed by the Business Goals Approach measure.

The role of Valence

Requirements Must was explained by Valence Obstruct and Requirements Won't was marginally explained by Valence Support. Thus, it might be that in the Business View, valence was not a moderator but a mediator, in between goals on the one hand and requirements on the other. To test whether valence should be conceived of as a mediator in the Business View, I followed the procedure suggested by Baron and Kenny (1986) for identifying mediating variables.

I ran a multiple linear regression analysis of Business Goals Approach and Valence Support on Valence Obstruct. Significant results were obtained, R^2= .57, R^2_{adj}= .51, $F_{(2,14)}$= 9.15, p= .003, indicating that mainly the correlated Business Goals Approach contributed to Valence Obstruct, standardized β= 79, t= 4.23, p= .001, $r_{partial}$= .75, r_{part}= .74.

Yet, if Valence Obstruct indeed was a mediator, then omitting it from the analysis should increase the predictive power of Business Goals Approach and Business Goals Avoid (Baron & Kenny, 1986). Therefore, I performed another regression analysis on Requirements Must with Business Goals Approach and Business Goals Avoid as the predictors. However, this analysis yielded insignificant results, R^2= .18, R^2_{adj}= .06, $F_{(2,14)}$= 1.53, p= .250.

I also tested whether the Valence Support variability could be explained by Business Goals Avoid and Valence Obstruct entered in the first step and Business Goals Approach in the second step. However, no significant effects were established (Hoorn & Kok, 2005, Tech. Rep. [CD]). In all, valence cannot be regarded as a mediator between requirements and goals in the Business View.

4.3.2.7 Discussion of the Relation between Management-viewpoint Requirements and Business Goals

H1 stated that goals to approach explain must requirements through the mediation of valence. Yet, the results of the regression of Valence Obstruct on Requirements Must in the Business View caused the first crack in my conviction. First, this analysis pointed out that Business Goals Approach had no significant explanatory power for the Must variability and that Valence Obstruct seemed an independent predictor instead of a mediator, explaining a large portion of the variance in the agreement to Requirements Must singlehandedly (standardized β= 1.03, $r_{partial}$= .71, r_{part}= .60).
Valence Obstruct was not a mediator between Approach and Must because Business Goals (either Avoid or Approach) could not explain the Requirements Must variability if Valence Obstruct was removed from the analysis. On the other hand, Business Goals Approach did make a significant contribution (standardized β= 79, $r_{partial}$= .75, r_{part}= .74) to Valence Obstruct. So much for H1.

H2 stated that goals to avoid predict won't requirements through the mediation of valence. However, I found that the opposite was the case. The multiple linear regression on Requirements Won't showed that Business Goals *Approach* (not Avoid) was the best predictor, explaining the variability in the level of agreement to Requirements Won't above and beyond the Valence Support measure (β= .69, $r_{partial}$= .81, r_{part}= .38). Moreover, Valence Support was not explained by Business Goals Approach (nor Avoid), which canceled out the possibility that Valence Support was a mediator. Valence Support probably served as a moderator.

I went back to the data files and questionnaire forms to see whether a labeling mistake was made. I also controlled extra for outliers, skewness, and other trouble in the data but nothing suspicious could be found (Hoorn & Kok, 2005, Tech. Rep. [CD]). Apparently, I was forced to take the reversal of H2 seriously. Here, the data told me that things stakeholders wanted to achieve with a system (e.g., to save money) should be accomplished by things the sys-

tem should not allow (e.g., planning based on availability of personnel instead of expected activity). This Won't-Supports-Approach triad probably indicated that officers said to their managers, for example, "You don't want informal hour registration but exactly this is what will help you save some money."

I thought this to be an intriguing possibility to explain change requirements. Figure 5 displays the results and prompted the following interpretation. These officers seemed to say to their managers: "The won't requirements are good for what you want to reach" (Won't Supports Approach). "The must requirements are bad for what you want to reach" (Must Obstructs Approach). "But both types of requirements are inconsequential to what you want to avoid" (Avoid had no significant explanatory power whatsoever).

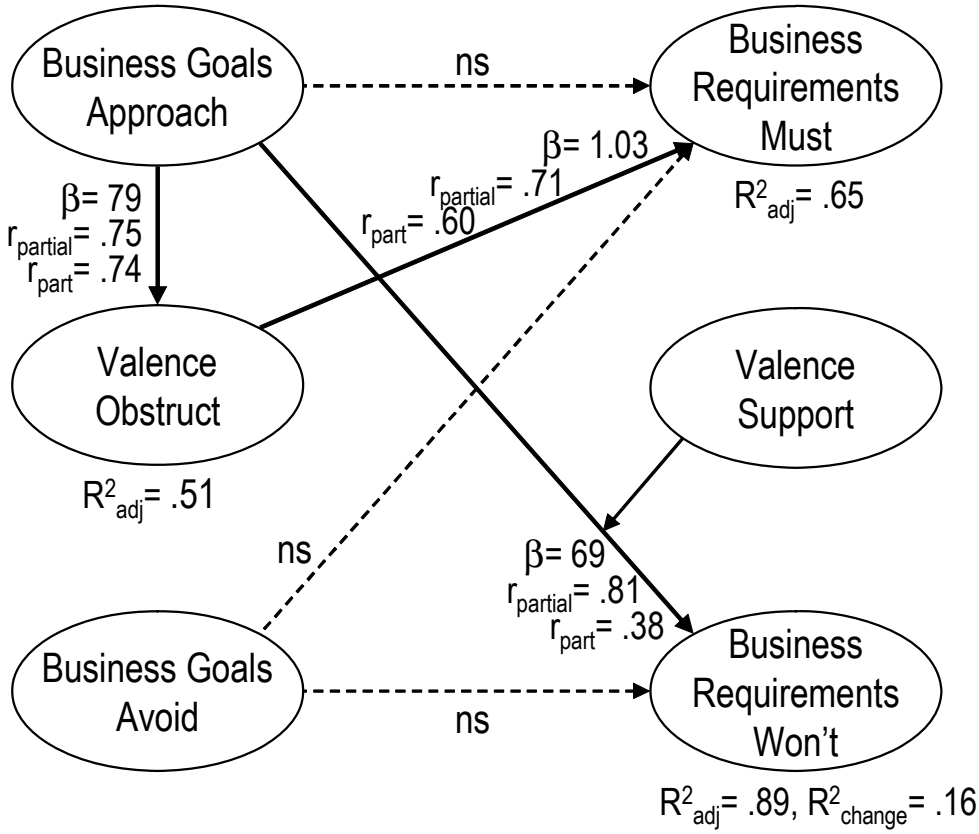


Figure 5: Goals to Approach explain the Won't Requirements.[8]

If stakeholders really thought that goals to approach were strongly related to won't requirements, one expects the mirror image of that structure between goals to avoid and must requirements. Figure 5 shows that this was not the case. The relation between Business Goals to Avoid and Requirements Must

[8] These are two sub models ((Approach→Obstruct, Obstruct→Must) and Avoid→Won't). Beta weights and correlations should not be compared.

was not established. Valence Obstruct seemed to take on that role a little bit but then in connection to Business Goals to Approach.

Given the mixed results, a number of replication studies were needed (Ohlsson & Runeson, 2002). The odd thing about Study 1 was that the management-viewpoint requirements were unaligned with the personal view of the officers. In the Business View, then, officers who personally disagreed to the business goals (Figure 3) had to take the business' position and judge business-related requirements. It might be that taking this to them somewhat unnatural position was responsible for the failure of H1 and H2 in the Business View.

4.3.3 CMS: Practitioner's Perspective

Evelien Kok works as functional analyst at the Concern Information Management Police.[9] She is responsible for the design standards of user interfaces and helps to develop and introduce the CMS nation wide. The reason why she as a practitioner turned to academic assistance is that she wanted to know whether the systems she developed could and should be tailored to the wishes and demands of the end-users; in particular, when such systems (i.e. the CMS) were more-or-less forcing particular work processes upon the personnel so to increase efficiency.

She convinced her project manager of the importance of the research because it would render information on how the CMS would be received and would help to anticipate the level of user acceptance. She wrote: "If a system does not fit the wishes and demands of the users, it will be hard to get it accepted, with all the consequences for system implementation and optimally using it." After a pitch to the management board with a number of sample items she got the OK.

The main problem this practitioner experienced in constructing the questionnaire items on the Stakeholders' Needs scale was with stating negative expectations about must requirements for goals to approach. For example:

< Digital incident reports (Requirements Must)> <hinder (Valence Obstruct)> < resolving urgent matters quickly (Goals Approach) >

Another difficulty was to formulate positive expectations about won't requirements for goals to avoid:

< Planning based on available personnel (Requirements Won't)> <prevents (Valence Support)> < that special operations run out of hand (Goals Avoid) >

Because such statements feel somewhat counter-intuitive, the practitioner felt the need to use linguistic negations (not, no, never) to make the items better understandable. In a pre-test, however, the items we constructed with such negations contributed poorly to the reliability of the scale and had to be re-

[9] www.cip.politie.nl

moved, she admitted. The good news was that after having read two qualitative research papers while gaining more experience with quantitative research through questionnaire design, the practitioner decided to add more quantitative research to her RE (personal e-mail communication, Oct. 27, 2003).

Later, she testified that our work made her aware of different and other pitfalls during requirements gathering and evaluation than she usually thought of. One of the clear changes in this practitioner's approach is that she more-or-less abandoned classic task analysis and relied less on ethnography so to work with persona's that are goal and viewpoint driven (personal e-mail communication, Aug. 3, 2005). Regarding the requirements-analysis rift, she replied "Interesting results. I can imagine that your analysis is right. The police has the obligation to make sure that civilians live by the rules and that for this reason the average copper is quite loyal as well [to the organization]. At least that is what I notice in my work environment" (personal e-mail communication, Sept. 6, 2005).

4.4 Study 2: Business Viewpoints on Requirements and Goals Aligned

In Study 2, I was anxious to evade the pitfall of Study 1, in which stakeholders expressed agreement to requirements from a viewpoint that was not their own. To make up for this peculiarity, I wanted managers to judge management-viewpoint requirements from a business perspective. I was invited by Business Mathematics and Computer Science (i.e. Sandra Pronk) to participate in the design of a logistic warehouse-management system (LWMS) at one of the provincial governments in our country. Apart from designing the system, I entered this project with four hypotheses in mind. Two were the classics

- (H1) Goals to approach predict agreement to the must requirements through the mediation of positive outcome expectancies (valence support)

- (H2) Goals to avoid predict (dis)agreement to won't requirements, mediated through negative valence (valence obstruct)

whereas the other two hypotheses were their competitors

- (H3) Goals to approach predict agreement to the won't requirements while being moderated by positive outcome expectancies (valence support)

- (H4) Goals to avoid predict agreement to must requirements, moderated by negative expectations (valence obstruct)

Although H4 was not derived from the empirical data directly, it was the logical counterpart of H3. The role of valence was also different from what I originally thought. In H1 and H2, I conceived of valence as a mediating vari-

able necessary to predict the level of agreement to requirements. In H3 and H4, I thought of valence in a moderating role, as a possible influence on agreement.

4.4.1 Method

4.4.1.1 Participants and Experimental Design

Managers (N= 18; 11 male, 7 female; age M= 46.4, SD= 10.9; years in service M= 14.4, SD= 11.7) from a provincial governmental institution in The Netherlands participated in a questionnaire study that concerned the (re)design of the LWMS. These participants ranged from various services, sectors, and functions within the organization.

The experimental design consisted of just one group working from one perspective, the Business View (within-subjects), while expressing their level of agreement to Must and Won't Requirements that could Support or Obstruct Goals to Approach or to Avoid.

4.4.1.2 System

The state of the warehouse management system at the time of measurement was a mainly manually and personally driven order and delivery system without intensive automation. Errors occurred regularly but were corrected effectively although not fast. (Re)designing this system was directed at higher efficiency, cost-effectiveness, and fewer behavioral rules while maintaining the current flexibility. The future system aimed at introducing Intranet and e-mail facilities to handle orders and deliveries while reducing the number of human transactions (Pronk, 2004, Tech. Rep. [CD]).

4.4.1.3 Procedure

As part of an internship with the said provincial government (Pronk, 2004), rapid ethnography (Jordan, 1996; Norman, 1998) in the early stages of design established a list of features of the current system, a list of requirements on the future system as well as a list of goals of the managers of the organization (not necessarily the same people who participated in the questionnaire study). Based upon these observations, a structured questionnaire, the LWMS RE*quest* (Appendix 4.1), of 64 items was created (in Dutch), divided into 5 blocks. Three blocks were created for the purposes of the IT practitioner who performed the internship, one block was created for hypothesis testing, and one block concerned socio-demographic information of the managers. The block for hypothesis testing was put in between the practitioner's blocks. The block of socio-demographic items was put in last. Items were pseudo-randomly distributed over blocks. Thirty-five participants were asked to print and fill out this paper-and-pencil questionnaire, which was sent to them over the e-mail. After a few reminders, eighteen questionnaires were completed and returned, which took about a fortnight.

Sample requirements in the LWMS case

- Must (Business)
  - Direct ordering at the warehouse
  - Direct access with personal computer
  - Inspecting status of order with personal computer
  - E-mail notification of order delivery
  - Reply e-mail for order acceptance/authorization
  - Warning e-mail that processing the order went wrong
- Won't (Business)
  - Know exactly where to place an order (which organizational unit handles what)
  - Follow the standard procedures (through several units)
  - Asking different people what the order status is
  - Order delivery without notification
  - Written autograph on paper receipt
  - Find out yourself if something went wrong with the order

Sample goals in the LWMS case

- Approach (Business)
  - Order-process control
  - Flexible procedures
  - Quick order handling
  - Accurate order handling
  - Proper planning
  - Work efficiently
  - Save money
- Avoid (Business)
  - Confused order process
  - Inflexible procedures
  - Slow order handling
  - Inaccurate order handling
  - Sloppy planning
  - Work inefficiently
  - Spend money

4.4.1.4 Measurements

Scale construction in the LWMS case

We created a Stakeholders' Needs scale (24 items plus 4 fillers) as explained in Section 4.3.1.4. It consisted of three times two sub scales: Requirements (Must vs. Won't), Valence (Support vs. Obstruct), and Business Goals (Approach vs. Avoid). Requirements were gathered during the internship. Must requirements covered aspects of automation and digitalization of operations whereas Won't requirements keyed manual aspects and human interference that was typical for the old system. Business Goals were divided into goals to Approach or goals to Avoid. Goals covered aspects of time efficiency, error reduction, and cost-effectiveness. Valence was operationalized as keying Support or Obstruction of goals by the specific requirement. An example of an item is "An e-mail warning that something is wrong with my order (a Must) enables (Support) working efficiently (Approach)." For more example items, see (Appendix 4.1). Items were followed by a 6-point rating scale (0= completely disagree, 5= completely agree). In addition, we investigated the agreement to the features of the Current System (BTM1 – Chapter 2) by means of Likert-type items (Appendix 4.1). Further, socio-demographic information was sampled, such as sex, age, service, sector, function, and number of years in function. Two staff members who were not involved in the actual test checked the items for readability and understandability.

4.4.2 Analysis and Results

After the completed questionnaires were returned, the data were entered in an SPSS 11.0 data matrix for statistical analysis. Details about the statistical procedures and intermediate results can be found in (Hoorn, 2004, Tech. Rep. [CD]).

4.4.2.1 Scale Analysis

I checked the reliability of the 12 items on each sub scale with Corrected Item-Total Correlations and Standardized Cronbach's alpha. The extent to which items were independent of other scales was verified with Pearson correlations. I followed the same procedures and criteria as described in Section 4.3.2.1.

The thus revised sub scales had 3 items each. Standardized Cronbach's alpha of four revised sub scales ranged from .61 to .78. However, Requirements Won't (.48) and Valence Obstruct (.50) were poor measurements and could not be improved. Results obtained with these sub scales should be taken with care and interpreted in the context of subsequent replication studies. For more details and the psychometric results of the Current System scale, see (Hoorn et al., 2005).

4.4.2.2 Agreement to LWMS Requirements and Goals

I checked to what extent the managers of the provincial institution agreed with the management-viewpoint requirements on the LWMS and whether they agreed with the business goals we obtained. We also inspected the effects of socio-demographic variables (e.g., age, sex, and function).

I treated the faceted scale of Stakeholders' Needs as a nested factorial design (within-subjects) of the 3-leveled factor Scales (Requirements vs. Valence vs. Business Goals) and the 2-leveled factor Item Type (Indicative vs. Contra-indicative). Six within-subjects (dependent) variables were calculated from the 3 items per sub scale: The grand mean average level of agreement to Requirements (Must vs. Won't) vs. Valence (Support vs. Obstruct) vs. Business Goals (Approach vs. Avoid). As a preliminary test, a One-Way MANOVA was run to check the potential effects of the fixed factors Service (4), Sector (7), and Sex (2) on the grand means of the 6 within-subjects (dependent) variables. The effects of Age (28-58) and Number of Years in Service (1-36) were controlled for by treating them as covariates. Function (14) was not analyzed because each function had but one or two managers. Multivariate tests according to Pillai showed that none of the fixed or covariate factors were significant ($.36 < F < 1.59$; $.479 \leq p \leq .700$) for either of the dependents.

In addition, the main test consisted of a 2*3 MANOVA of Item Type (Indicative vs. Contra-indicative) (within-subjects) and Scales (Requirements vs. Valence vs. Goals) (within-subjects) on the grand mean average agreement to the 6 sub scales. Results can be found in Figure 6.

A significant interaction occurred between Item Type (Indicative vs. Contra-indicative) and Scales (Requirements vs. Valence vs. Business Goals) (Pillai's Trace= .51, $F_{(2,16)}$= 8.40, p= .003). Parameter estimates showed that Indicative items of Requirements (M_{Must}= 2.41, SD= .98) evoked higher levels of agreement than Contra-indicative items ($M_{Won't}$= 1.80, SD= 1.09), which may be expected. This difference was larger, however, for Business Goals. Indicative items of Business Goals ($M_{Approach}$= 3.67, SD= 1.14) evoked the highest level of agreement in this study, more than contra-indicative items (M_{Avoid}= 2.50, SD= .96) (parameter coefficient= -.56, t= -4.04, p= .001, η_p^2= .49).

The Indicative items of Valence ($M_{Support}$= 2.19), surprisingly, elicited *lower* levels of agreement than the Contra-indicative items ($M_{Obstruct}$= 2.78). As mentioned in the previous paragraph, the opposite happened for Business Goals (parameter coefficient= -1.76, t= -3.25, p= .005, η_p^2= .38).

The third contrast was marginally significant according to Bonferroni (α= .05/3≈ .017) and should be considered a trend. The level of agreement to Indicative and Contra-indicative items in Requirements had an inverse pattern as compared to Valence (parameter coefficient= 1.20, t= 2.51, p= .022, η_p^2= .27).

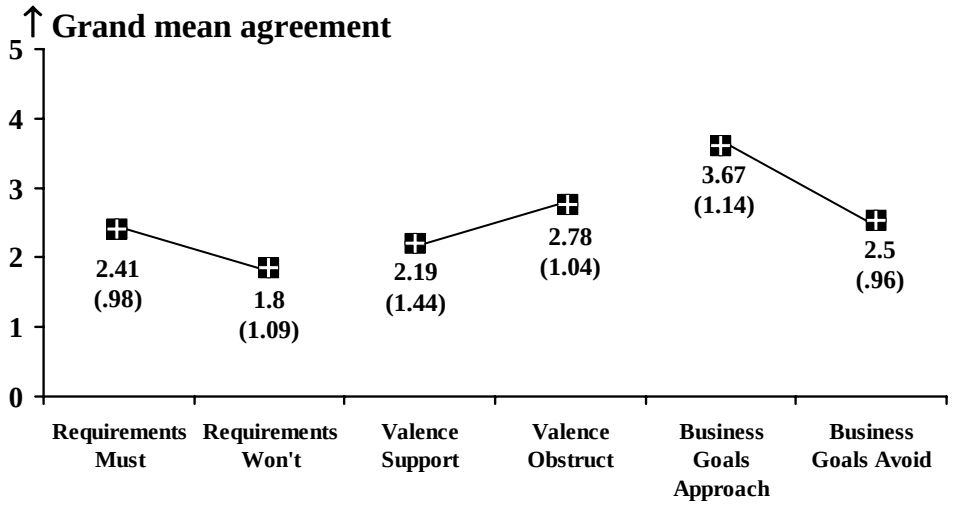


Figure 6: Grand mean average agreement to the 6 sub scales of Stakeholders' Needs (N= 18). Standard deviations are in parentheses.

These interactions were sustained by a significant main effect of Scales (Pillai's Trace= .44, $F_{(2,16)}$= 6.40, p= .009), which was mainly based on the contrast between Requirements and Business Goals (parameter coefficient= -1.96, t= -3.57, p= .002, η_p^2= .44). The difference between Valence and Business Goals was much smaller and only marginally significant (parameter coefficient= -1.20, t= -2.34, p= .032, η_p^2= .24) according to Bonferroni (.05/3≈ .017). In other words, the strongest interactions and main effects were produced by Business Goals in combination with Requirements, whereas the weaker interactions and main effects were generated by Valence in combination with Business Goals.

4.4.2.3 Discussion of the Effects on Agreement in the LWMS Case

I aligned the management-viewpoint requirements with the goals in the Business View and found that the managers agreed more to the Business Goals than to the Requirements. This may be expected because requirements are but one instantiation of business goals. Other requirements could do just as well or perhaps even better. The ratios went in the expected direction, though. Business goals to Approach raised more agreement than goals to Avoid and Must requirements more than Won't requirements. Although we gathered both requirements and goals from colleague-managers, the managers in the test sample felt that on the whole the requirements would frustrate the business goals ($M_{Obstruct}$= 2.78) rather than sustain them ($M_{Support}$= 2.19). Whether this controversy among the managers affected the relations between requirements and goals is inspected next.

4.4.2.4 LWMS: Management-viewpoint Requirements in Relation to Business Goals

I formulated two competing sets of hypotheses. H1 and H2 advocated the common sense approach to requirements analysis. Must requirements are fed by goals to approach and mediated by expectations of support (H1). Won't requirements are fed by goals to avoid, mediated by expectations of obstruction (H2).

H3 and H4 countered these assumptions, referring to the Business View data of the CMS police case in Study 1. H3 assumed that goals to approach directly explain won't requirements and is moderated by expectations of support. H4 assumed that goals to avoid directly explain must requirements, moderated by expectations of obstruction.

Explaining Requirements Must

Requirements Must served as the dependent variable in a multiple regression (method Enter) with three ordered sets of predictors. Requirements Won't was entered in the first step, Valence Obstruct and Business Goals Avoid in the second step, and Valence Support and Business Goals Approach in the third (Hoorn, 2004, Tech. Rep. [CD]).

Business Goals Avoid and Valence Obstruct together accounted for a significant quantity of the Requirements Must variability, R^2= .93, R^2_{adj}= .90, $F_{(5,12)}$= 30.30, p= .000. Business Goals Approach and Valence Support did not significantly increment the percent of explained variance of Requirements Must, R^2_{change}= .01, $F_{(2,10)}$= .33, p= .728. I also assessed the relative importance of Business Goals Avoid and Valence Obstruct in predicting Requirements Must. Business Goals Avoid was most strongly related to Requirements Must (standardized β= -.97, t= -9.48, p= .000). Supporting this conclusion is the height of the standardized Beta coefficient and the strength of the correlation between Business Goals Avoid and Requirements Must, partialling out the effects of all other predictors ($r_{partial}$= -.94, r_{part}= -.74). Valence Obstruct offered little or no additional predictive power beyond that contributed by the Business Goals Avoid measure. The Current System scale merely rendered insignificant effects (Hoorn et al., 2005).

Explaining Requirements Won't

Business Goals Approach and Valence Support accounted for a significant amount of the Requirements Won't variability, R^2= .79, R^2_{adj}= .70, $F_{(5,12)}$= 9.01, p= .001. Business Goals Avoid and Valence Obstruct did not increase the percent of explained variance of Requirements Won't, R^2_{change}= .07, $F_{(2,10)}$= 2.28, p= .153. I also assessed the relative importance of Business Goals Approach and Valence Support in predicting Requirements Won't. Business Goals Approach was most strongly related to Requirements Won't, standardized β= -.96, t= -5.31, p= .000. Supporting this conclusion is the height of the standardized Beta coefficient and the strength of the correlation

between Business Goals Approach and Requirements Won't, partialling out the effects of all other predictors ($r_{partial}$= -.84, r_{part}= -.70). Valence Support offered little or no additional predictive power beyond that contributed by the Business Goals Approach measure. The Current System scale merely yielded insignificant results (Hoorn et al., 2005).

The role of Valence

H1 further predicted that Business Goals Approach measure explains Valence Support. H2 predicted that Business Goals Avoid measure explains Valence Obstruct. However, no significant results were obtained in the respective regression analyses (Hoorn, 2004, Tech. Rep. [CD]).

4.4.2.5 Discussion of the Relation between Management-viewpoint Requirements and Business Goals

H1 expected that requirements the system must meet are explained by a positive outcome valence of the proposed features towards goals the stakeholder wants to achieve in his or her work. The opposite was the case, however, confirming H4. Business Goals Avoid significantly accounted for a large portion of the variability in agreement to Requirements Must (standardized β= -.97, $r_{partial}$= -.94, r_{part}= -.74). Observe that this theoretical possibility was not realized in the CMS case (Study 1) but hypothesized purely on the basis of symmetry with H3.

The mirror image of the above structure was found for the requirements the system won't have. H2 anticipated that what the system won't have is predicted by a negative outcome valence of the proposed features towards states and situations the stakeholder wants to avoid in his or her work. Again the reverse happened, confirming H3. Business Goals Approach significantly accounted for a substantial amount of variability in agreement to Requirements Won't (standardized β= -.96, $r_{partial}$= -.84, r_{part}= -.70), despite the latter measure's poor quality.

As another matter, H1 and H2 assumed that valence was a mediator between agreement to requirements and goals. This was not demonstrated by the regression results in Study 2, however. The relative importance of Business Goals Avoid to Requirements Must was significantly higher than for all other predictors, including Valence ($r_{partial}$= -.94, r_{part}= -.74). Likewise, the relative importance of Business Goals Approach to Requirements Won't also was significantly higher than for all other predictors, including Valence ($r_{partial}$= -.84, r_{part}= -.70).

Valence moderated the relational strength between goals and requirements. On the one hand, the MANOVA in Section 4.4.2.2 showed that Valence was involved in a significant interaction with Business Goals on agreement. On the other hand, Valence had no significant main effect according to Bonferroni. Additional multiple regressions indicated that Business Goals Approach did not significantly predict Valence Support and that Business Goals

Avoid did not significantly predict Valence Obstruct. Therefore, valence should be regarded a moderating rather than a mediating variable.

Agreement or disagreement to the features of the Current System (BTM1 – Chapter 2) lacked explanatory power for agreement to must and won't requirements (BTM2). In sum, Study 2 is a clear-cut rejection of H1 and H2 in favor of H3 and H4 (Figure 7).

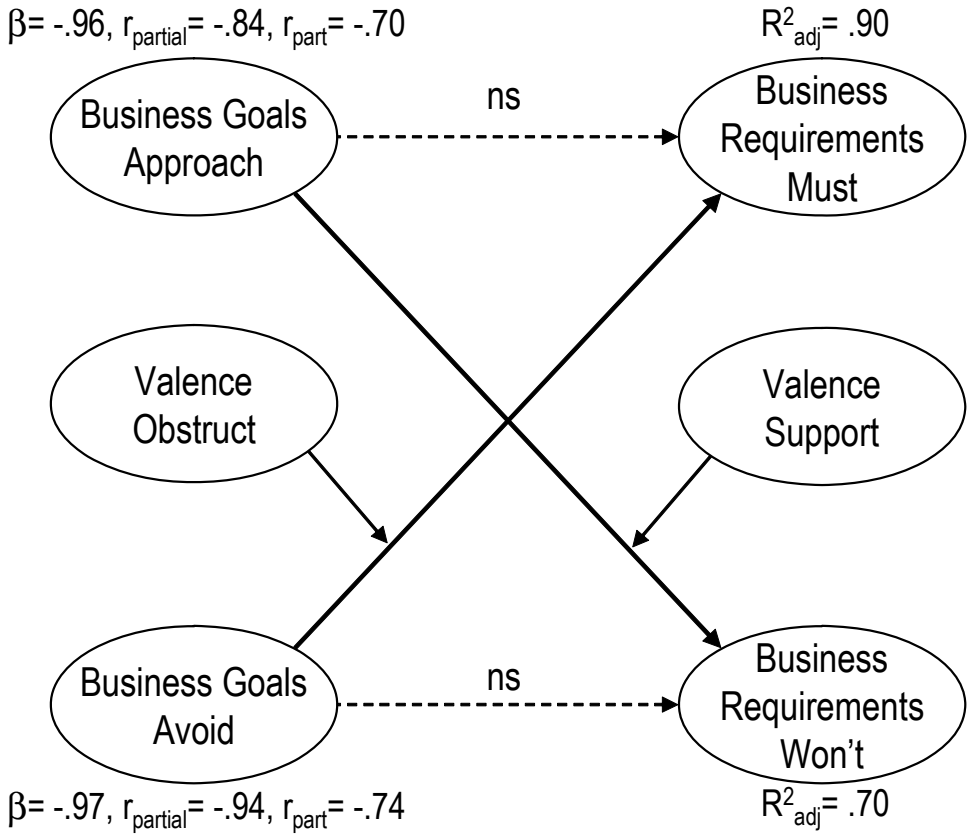


Figure 7: Goals to Approach explain variability in the Won't Requirements, Goals to Avoid in the Must Requirements.[10]

4.4.3 LWMS: Practitioner's Perspective

Sandra Pronk performed an internship with the provincial government to design and develop the LWMS. Her official assignment was to "investigate, restructure, and reorganize the warehouse space and to redefine the functions and procedures of the current warehouse." A high level of automation was demanded for better warehouse management ("from classic store to automated warehouse"). Another task was to make the stakeholders aware that this process was relevant and that a logistic update was necessary. The original

[10] The Beta weights and correlations of the two sub models (Avoid→Must and Approach→Won't) should not be compared.

assignment merely wanted an inventory of requirements on the basis of which the practitioner could write an advice to purchase commercial software for managing the warehouse processes. The choice for a software package was supposed to reckon with the human side of things (i.e. accessibility), other software systems that ran in the organization, and the type of goods that should be transported (personal e-mail correspondence, March 17, 2004 –May 18, 2004).

What the practitioner expected as added value of running the LWMS RE-*quest* was the scientific validation of requirements and its focus on the human aspects, so that she could write a better software advice. The problems she envisioned were the project's timeline, the usefulness of the conclusions, and 'selling the extra effort' to her managers. Some of the concerns she reported were "When the analysis is completed, can I directly use the conclusions? I need to write my advice and deliver a system design. If I already made a choice and the conclusions come late I need to start searching all over again. Both my boss and I think that running the questionnaire should also benefit our system design and not merely answer a scientific question" (ibid.).

The process manager that supervised the practitioner was eager to get involved in designing the questionnaire. This way, he made sure that the results were beneficial to the organization and that the list of questions did not become too long. The process manager also expected that not many stakeholders would be willing to participate and that the group of relevant stakeholders was small anyway (no more than 30). One of the conditions of running the questionnaire was that the organization and the results were kept anonymous. One person in the organization with some experience in survey design had to be convinced about the use of the questionnaire and why it featured options that could not be implemented or were undesirable anyway (the won't requirements). She functioned as a self-acclaimed advisor of the practitioner and the process manager (ibid.).

Yet, despite these concerns and challenges, we were able to convince the managers about the questionnaire's usefulness. One of the arguments was that running the questionnaire would make people aware that system redesign was necessary and that the opinion of the stakeholders was considered important. After the process manager scrutinized the final version, he gave us clearance to run the LWMS RE*quest* after which I analyzed the data quick enough to have the results sustain the advice. In her final report, this practitioner stated that this academic approach was a welcome addition to standard engineering practices and helped to evaluate the MuSCoW list (Pronk, 2004, pp. 26-27, pp. 60-61). After finishing the internship, she was offered a steady job to implement and improve the LWMS, which she accepted.

4.5 Study 3: Personal Viewpoints on Requirements and Goals Aligned

The alternative hypotheses H3 and H4 were derived from the CMS data in the Business View (Study 1) and confirmed by the LWMS case in the Business View (Study 2). My next attempt, therefore, was to discover the surprising crossover between requirements and goals in the Personal View. This did not work out in Study 1, probably because the management-viewpoint requirements were unaligned with the personal viewpoint of the stakeholders. The possibility of stumbling into the requirements-analysis rift (Chapter 3) was avoided, then, by sampling personal requirements in alignment with personal goals of the stakeholders. I did so for end-users of Commercial Off-the-Shelf (COTS) systems (PCs).

Because of the results of Study 2, I was now able to formulate precise predictions about how the new model of requirements change should look like in real life cases; the relations that should be confirmed or rejected to speak of the goals-to-requirements chiasm (χ-effect). The third study into requirements on COTS systems was set up to verify eight predictions as derived from H3 and H4. These predictions formed a sort of check list to find out how many relations that the χ-effect predicted could be established in a business case.

The first set of six predictions can be regarded as direct confirmation of the χ-effect. The second set of two predictions pertains to the status of valence. They specify a different position of valence in the chiasm than expected. If these two come true, they are 'in line' with the χ-effect. All other relations among variables should be excluded and count as counter-evidence.

Confirmation of the χ-effect

(P1) Won't requirements depend on goals to approach (and some random noise factors) (Approach→Won't)

(P2) Must requirements depend on goals to avoid (and some random noise factors) (Avoid→Must)

(P3) The dependency of won't requirements on goals to avoid is insignificant ¬(Avoid→Won't)

(P4) The dependency of must requirements on goals to approach is insignificant ¬(Approach→Must)

(P5) Positive expectations moderate the relation between won't requirements and goals to approach

$$(\text{Approach} \overset{\text{Support}}{\pm\downarrow} \text{Won't})$$

(P6) Negative expectations moderate the relation between must requirements and goals to avoid

$$(\text{Avoid} \overset{\text{Obstruct}}{\downarrow} \text{Must})$$

In Line with the χ-effect

(P7) Positive expectations mediate the relation between won't requirements and goals to approach (Approach→Support→Won't)

(P8) Negative expectations mediate the relation between must requirements and goals to avoid (Avoid→Obstruct→Must)

Refutation of the χ-effect

All other relationships.

4.5.1 Method

4.5.1.1 Participants and Experimental Design

Experts in computer-human interaction (CHI) (N= 14; 12 male, 2 female; age 22-43, M= 29.7, SD= 6.22) participated in a questionnaire study that concerned assembling a computer system with off-the-shelf products. These CHI experts gathered in a business meeting of two different knowledge-intensive organizations except for one, who was a visitor from a third organization. Together, these people represented six different nationalities (Holland, Israel, Poland, India, Romania, and Russia).

Stakeholders' Needs was a nested within-subjects factor of Requirements (Must vs. Won't) vs. Valence (Support vs. Obstruct) vs. Personal Goals (Approach vs. Avoid). Requirements Must consisted of a selection of system features taken from standard PCs as offered in state-of-the-art computer magazines. Requirements Won't consisted of outmoded system features such as plain DOS machines and high-radiation cathode ray tubes. Requirements were explicitly connected to personal goals with a Valence attached to it (positive outcome expectancies of using the requirements vs. negative expectancies). Personal Goals pertained to work-related personal concerns such as individual health, personal budget, and work pace.

4.5.1.2 Procedure

A structured Requirements Engineering questionnaire for experts of Computer-Human Interaction was assembled, the CHI RE*quest* (cf. Section 4.3.1.4). It consisted of 24 requirements statements and seven socio-demographic items (in English) (Appendix 4.2). Items were pseudo-randomly distributed.

During a business meeting between the two organizations (and one person listening in), the CHI experts were asked to fill out the paper-and-pencil questionnaire after a short introduction. The introduction emphasized that the requirements should be judged from a perspective of personal goals (e.g., available budget, personal work pace, personal health, and individual effort). Completing the questionnaire took between 10 and 15 minutes.

Sample requirements in the COTS PC case

- Must (Personal)
  - Anti virus software
  - 5 years pickup, repair, and return guarantee
  - Linux operating system
  - Latest AMD Athlon 64 processor
  - 350 GB hard disk
  - Mouse device
  - Internet connection
  - Well-designed interface
  - 63” Wide Screen Plasma Monitor
  - Outstanding firewall
  - TFT monitor
- Won’t (Personal)
  - Cathode ray tube monitor
  - Windows ’95 operating system
  - Outdated browser software
  - 1 GB hard disk
  - Green-on-black display screen
  - Second hand DOS machine
  - Stand-alone computer
  - 5¼” floppy drive
  - Working from the prompt only
  - 486 DX processor

Sample goals in the COTS PC case

- Approach (Personal)
  - Cheap buy / save cash / no fees
  - Work quickly
  - Nice look-and-feel
  - International access / communication
  - Store all files

- System usability
- System stability
- Good health

- Avoid (Personal)
  - Monthly license fee
  - System instability
  - Delete most files
  - Repetitive strain injury
  - High costs
  - Working in isolation
  - Work slowly
  - Hacker attack
  - Damage to eyes

4.5.1.3 Measurements

Scale construction in the COTS PCs case

The faceted scale of Stakeholders' Needs had 24 items, which were scored for agreement using 6-point rating scales (0= completely disagree, 5= completely agree). Stakeholders' Needs consisted of requirements statements that systematically connected personal goals to requirements, while putting positive or negative valence to them. Example items are "I want the latest AMD Athlon 64 processor so that I can work quickly" (Must-Supports-Approach), "To get a 63 inch Wide Screen Plasma Monitor I am willing to drain my budget" (Must-Obstructs-Approach).

The items were evaluated by ICT professionals (not the actual test group) for readability, wording, and whether the contents made sense to people working in the field. After the necessary repair work, items were considered fit for the main test.

4.5.2 Analysis and Results

Data were entered in an SPSS 11.0 data matrix for statistical analysis. For a detailed analysis, see (Hoorn, 2005a, Tech. Rep. [CD]).

4.5.2.1 Scale Analysis

In doing scale analysis, I followed the same procedures and criteria as described in Section 4.3.2.1. I did reliability assessments and item selection of the 12-item sub scales with Corrected Item-Total Correlations and Standardized Cronbach's alpha. I calculated Pearson correlations to evaluate in how far items correlated with other scales than their own.

The revised sub scales each had 3 items. On the whole, scale reliabilities were acceptable (.65) to reasonable (.77). For more details, see (Hoorn, 2005a, Tech. Rep. [CD]).

4.5.2.2 Agreement to COTS PCs Requirements and Goals

A 2*3 MANOVA of 2 (Item Type: Indicative vs. Contra-indicative) (within-subjects) * 3 (Scales: Requirements vs. Valence vs. Personal Goals) was run on the grand mean level of agreement to requirements statements (Figure 8). For the complete analysis, see (Hoorn, 2005a, Tech. Rep. [CD]).

The most important result was the significant interaction between Item Type (Indicative vs. Contra-indicative) and Scales (Requirements vs. Valence vs. Personal Goals) (Pillai's Trace= .78, $F_{(2,12)}$= 21.06, p= .000). The interaction underscored the general pattern that sub scales consisting of Indicative items (M_{Must}= 3.68, $M_{Support}$= 2.45, $M_{Approach}$= 1.41) raised higher levels of agreement than those with Contra-indicative items ($M_{Won't}$= .21, $M_{Obstruct}$= 1.79, M_{Avoid}= 1.10). For Requirements Must (M= 3.68, SD= .93) vs. Requirements Won't (M= .21, SD= .32) the difference was the largest. This was manifested in a strong significant contrast between Requirements on the one hand and Valence on the other (parameter coefficient= 2.80, t= 5.78, p< .0001, η_p^2= .72) as well as between Requirements and Personal Goals (parameter coefficient= 3.15, t= 6.69, p < .0001, η_p^2= .78). Yet, the Indicative and Contra-indicative items of Valence and Personal Goals differed in a similar way, making the contrast within the interaction insignificant (parameter coefficient= .36, t= 1.13, p= .281). Instead, the significant main effect of Scales (Pillai's Trace= .65, $F_{(2,12)}$= 11.09, p= .002) showed that on the whole, Valence (M= 2.12) evoked higher levels of agreement than Personal Goals (M= 1.26) did (parameter coefficient= 1.74, t= 4.29, p< .001, η_p^2= .59).

The entire analysis was repeated with Sex (2), Function (3), Organization (3), Department (3), Section (3), and Country (6) as fixed factors and controlled for Age as covariate. However, none of the effects of these background variables were significant (Hoorn, 2005a, Tech. Rep. [CD]).

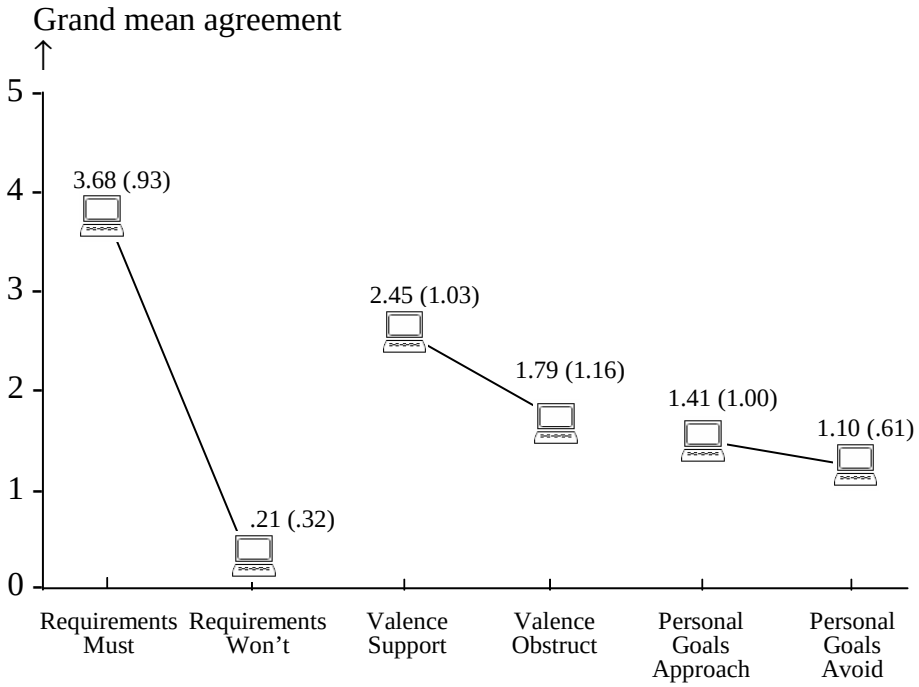


Figure 8: Grand mean average level of agreement to Stakeholders' Needs (Must, Won't, Support, Obstruct, Approach, Avoid). Standard deviations are in parentheses (N= 14).

4.5.2.3 Discussion of the Effects on Agreement in the COTS PCs Case

The significant interaction between Item Type and Scales underscored that the pattern of scores went into the expected direction. The CHI experts expressed more agreement to Must than Won't, to Support than Obstruct, and more to Approach than to Avoid.

4.5.2.4 COTS PCs: Personal-viewpoint Requirements in Relation to Personal Goals

Next, I confronted the predictions P1 to P8 with the data obtained for the COTS PCs. I performed a number of multiple linear regressions to determine the predictive power of Personal Goals Avoid for Must Requirements and Personal Goals Approach for Won't Requirements.

Explaining Requirements Must

A multiple linear regression analysis (method Enter) was performed in which Requirements Must served as the dependent variable and Personal Goals Avoid and Personal Goals Approach as the predictors (Hoorn, 2005a, Tech. Rep. [CD]). Personal Goals Avoid accounted for a significant amount of the Requirements Must variability, R^2= .77, R^2_{adj}= .73, $F_{(2,11)}$= 18.28, p= .000. Sustaining this conclusion is the height of the standardized Beta coefficient, standardized β= .86, t= 5.92, p= .000, and the strong correlation ($r_{partial}$= .87,

r_{part}= .86) between Personal Goals Avoid and Requirements Must partialling out the effects of Personal Goals Approach.

In a second multiple linear regression analysis (method Enter), Requirements Must was the dependent variable with three ordered sets of predictors. Requirements Won't was entered in the first step, Valence Support and Valence Obstruct in the second step, and Personal Goals Approach and Personal Goals Avoid in the third.

The second set of predictors, Valence Support and Valence Obstruct, accounted for a significant amount of the Requirements Must variability, R^2= .84, R^2_{adj}= .79, $F_{(3,10)}$= 17.74, p= .000. The third set, Personal Goals Approach and Personal Goals Avoid, did not increment the percent of explained variance of Requirements Must, R^2_{change}= .06, $F_{(2,8)}$= 2.49, p= .145. On the basis of correlation-regression analyses, the relative importance of Valence Support and Valence Obstruct in predicting Requirements Must was assessed. Valence Obstruct was most strongly related to Requirements Must, standardized β= .90, t= 6.21, p= .000, ($r_{partial}$= .89, r_{part}= .78). Valence Support offered little or no additional predictive power beyond that contributed by the Valence Obstruct measure.

To check whether Valence Obstruct mediated the relation between Personal Goals Avoid and Requirements Must, a third regression was performed in which Valence Obstruct served as the dependent variable and Personal Goals Avoid and Personal Goals Approach served as the predictors. Personal Goals Avoid accounted for a significant amount of the Valence Obstruct variability, R^2= .87, R^2_{adj}= .85, $F_{(2,11)}$= 38.35, p= .000, standardized β= .94, t= 8.75, p= .000, $r_{partial}$= .94, r_{part}= .93. Personal Goals Approach had no additional predictive power beyond Personal Goals Avoid.

Explaining Requirements Won't

A multiple linear regression analysis (method Enter) was performed in which Requirements Won't served as the dependent variable with three ordered sets of predictors. Requirements Must was entered in the first step, Valence Support and Valence Obstruct in the second step, and Personal Goals Approach and Personal Goals Avoid in the third.

The third set of predictors, Personal Goals Approach and Personal Goals Avoid, accounted for a significant proportion of the Requirements Won't variability, R^2= .79, R^2_{adj}= .65, $F_{(5,8)}$= 5.84, p= .015. It seemed that Personal Goals Approach was most strongly related to Requirements Won't, standardized β= .75, t= 4.11, p= .003. The correlation between Personal Goals Approach and Requirements Won't was also high after partialling out the effects of Personal Goals Avoid ($r_{partial}$= .82, r_{part}= .67). Personal Goals Avoid offered little or no additional predictive power beyond that contributed by the Personal Goals Approach measure.

In a second regression, Valence Support served as the dependent variable and Personal Goals Avoid and Personal Goals Approach as the predictors.

Both Personal Goals Approach and Personal Goals Avoid showed little or no predictive power, $R^2 = .13$, $R^2_{adj} = -.03$, $F < 1$.

4.5.2.5 Discussion of the Relation between Personal-viewpoint Requirements and Personal Goals

I confronted the predictions P1 to P8 with the results of the COTS PCs study and drew the following conclusions.

(P1) (Approach→Won't) was confirmed. Personal Goals Approach independently predicted Requirements Won't (standardized $\beta = .75$, $r_{partial} = .82$, $r_{part} = .67$).

(P2) (Avoid→Must) was confirmed. Personal Goals Avoid independently predicted Requirements Must (standardized $\beta = .86$, $r_{partial} = .87$, $r_{part} = .86$).

(P3) ¬(Avoid→Won't) was confirmed. Personal Goals Avoid offered little or no additional predictive power to the Won't variability beyond that contributed by the Personal Goals Approach measure.

(P4) ¬(Approach→Must) was confirmed. Personal Goals Approach offered little or no additional predictive power to the Must variability beyond that contributed by Personal Goals Avoid.

(P5) (Approach $\overset{\text{Support}}{\downarrow}$ Won't) was confirmed. Valence Support served as a moderator of the relation between Personal Goals Approach and Requirements Won't because it did not have significant explanatory power for Requirements Won't but did have significant effects on level of agreement to the requirements statements in the MANOVA.

(P6) (Avoid $\overset{\text{Obstruct}}{\downarrow}$ Must) was rejected (see P8).

(P7) (Approach→Support→Won't) was rejected (see P5).

(P8) (Avoid→Obstruct→Must) was confirmed. Valence Obstruct served as a mediator (standardized $\beta = .90$, $r_{partial} = .89$, $r_{part} = .78$), annihilating the predictive power of Personal Goals Avoid on Requirements Must. Yet, in its turn, Personal Goals Avoid explained Valence Obstruct (standardized $\beta = .94$, $r_{partial} = .94$, $r_{part} = .93$). Taken in unison, this constellation is typical for a mediating variable (Baron & Kenny, 1986).

All the predictions derived from the goals-to-requirements chiasm were confirmed or rejected as expected, except for P6. Although not a straightforward confirmation, Valence Obstruct as a mediator (P8) does not harm the basic principle of the χ-effect that variability in agreement to positive or negative requirements is predicted by goals of opposite polarity (Figure 9).

As a possible threat to the construct validity of the scales in the COTS PC case, one might emphasize that the Won't Requirements scale was quite extreme (e.g., green on black cathode ray tubes and 5¼" floppies). Indeed, this resulted in strong differences between the levels of mean agreement with Must and Won't Requirements (Figure 8), which impact on the effect size. However, not the magnitude but the *direction* of effects was predicted and confirmed. In this sense, the intended manipulation was quite successful. More importantly, the relationships between requirements and goals established through regression analyses were similar to the LWMS case, which had less extreme won't requirements. Therefore, if there is a threat to construct validity at all in the COTS PC case, it cannot be too severe.

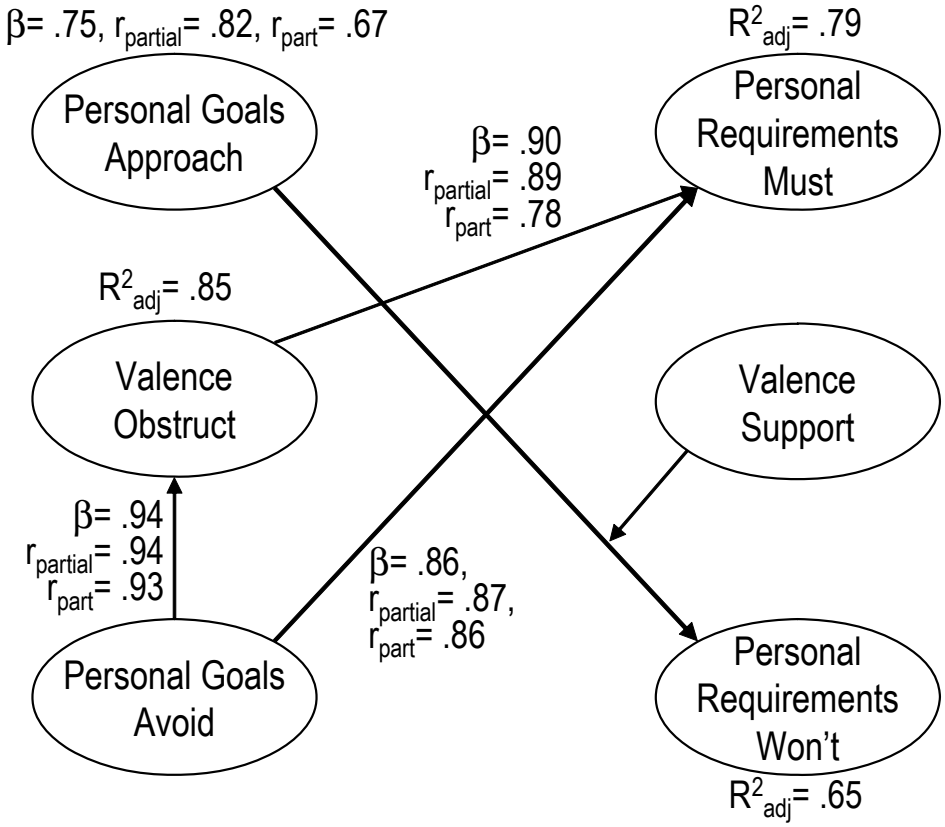


Figure 9: Goals to Approach explain variability in the Won't Requirements, Goals to Avoid in the Must Requirements. Negative expectations (Valence Obstruct) served as mediator.[11]

[11] The Beta weights and correlations of the two sub models (Avoid→Obstruct→Must and Approach→Won't) should not be compared.

4.6 Study 4: Personal Viewpoints on Requirements and Goals Revisited

Satisfaction is hard to find. On the one hand, Study 3 on the COTS systems showed that if the views on requirements and personal goals were aligned, the χ-effect occurred. If not, the requirements-analysis rift[12] caused the straight relationships as demonstrated in Study 1. On the other hand, one could counter that from a personal stakeholders' viewpoint, the game ended in a draw.

Moreover, the police officers in Study 1 were almost novices to the CMS, whereas the stakeholders of the COTS PCs (Study 3) were expert users. Perhaps the χ-effect was the result of a more sophisticated approach to requirements of experts as compared to the simpler approach of novices.

Furthermore, the moderator Valence Obstruct in Study 3 did not harm but also did not confirm the predictions. Due to reasons of working in real life environments, moreover, the measurements thus far were tested for psychometric quality post hoc.

In view of these considerations, I was more than happy to participate in the evaluation of a tactile mouse device and related software (the VTPlayer), produced for blind and low-vision users by the Israeli company VirTouch.[13] This gave me the opportunity to replicate the goals-to-requirements chiasm one more time, without the said restrictions.

4.6.1 Method

4.6.1.1 Participants and Experimental Design

Together with my master student Jelle van den Berg, I tested the VTPlayer with blind pupils (N= 15; 8 male, 7 female; age M= 15.07, SD= 1.03) from two colleges specialized in blind education: Bartimeus Foundation (Zeist, The Netherlands) and VISIO (Amsterdam, The Netherlands).

The experimental (within-subjects) design had one group of blind pupils working from a Personal View. They scored their level of agreement to Must and Won't Requirements that could Support or Obstruct Personal Goals to Approach or to Avoid.

4.6.1.2 System

The Israeli company VirTouch creates hardware and software that enables blind computer users access to interactive tactile graphics. Their flagship product is the patented VTPlayer tactile mouse. This device is a standard form-factor USB optical mouse with four buttons, two on each of the left and right sides, with the addition of two 4x4 pin tactile pads on the top of the mouse. Currently VirTouch offers a Learn & Play software series aimed at the K-12 market. Through the combined use of tactile, audio, and motion

[12] Stakeholders agree more-or-less to management-viewpoint requirements when they do not realize how those affect their personal goals.
[13] http://www.virtouch2.com/

feedback, this software teaches blind children hand, (mental) eye, and spatial coordination skills similar to those required for playing computer games designed for sighted children. Note that these coordination skills are critical for the effective day-to-day functioning of the blind, and not just for playing games. Typically it is quite difficult to teach these coordination skills to visually handicapped people and the VTPlayer is meant as a tool for teaching these critical skills. VirTouch also provides software games for learning Braille, which significantly decrease the Braille learning curve.

The VTPlayer is a haptic pointing device shaped like a regular but somewhat bigger computer mouse. The two Braille pads with 4x4 pins are mounted on top of the mouse, on the location where the index and middle finger are put. Mouse buttons are placed on both sides of the mouse to make it usable for both left- and right-handed users.

The VTPlayer outputs contrasts in a small square (4x4 pixels) under the cursor. Light colors (white) result in flat pins and dark colors (black) are outputted in rising pins. This process facilitates haptic exploration of imagery as if it were regular tactile maps. These drawings in combination with the VTPlayer make up the system. Like this, pupils can, for example, learn geography (which is hard for blind people) and play the games.

4.6.1.3 Procedure

To improve our measurements, we tested the items in a questionnaire, called the VTPlayer PRE*quest*. After psychometric analysis, we repaired the items when necessary and ran the VTPlayer RE*quest* with a sample of blind pupils.

VTPlayer PREquest

The participants of the VTPlayer PRE*quest* (N= 21, 15 male, 6 female; age M= 16.1, SD= 0.68) were sighted college students of the same educational level as the targeted VTPlayer RE*quest* group. We had to employ sighted students because of the relatively small number of blind pupils we could use for the main test.

After studying the literature on blind education, ethnographic research in blind schools, and additional interviews with teachers of blind pupils, a list of personal requirements for tactile geography maps was developed as well as a list of personal study goals. A questionnaire of 96 items (12 items for each of the eight item categories) was developed (in Dutch) combining requirements and goals. Items were structured according to Section 4.3.1.4. They were randomized and divided into 8 blocks. These items were preceded by 10 items on the relevance of personal goals. The items were checked for readability by a language teacher of the sighted pupils.

Prior to the test, the sighted pupils received a small demonstration of the VTPlayer, showing the working of the device while exploring a specially designed audio-tactile map of Europe. The pupils were asked to close their eyes and imagine being blind while acquiring some hands-on experience with the

VTPlayer. They completed the paper-and-pencil VTPlayer PRE*quest* thereafter, which took about 30 minutes.

Reliability analysis according to the criteria explained in Section 4.3.2.1 prompted the necessary adjustments. Psychometrically bad performing items were dismissed from the list, whereas moderately good items were repaired. The items were again checked for readability, this time by the principal of Bartimeus.

VTPlayer REquest

The VTPlayer RE*quest* was run at the respective colleges. The blind pupils were introduced to the VTPlayer by demonstrating the practical use of the mouse, how to explore imagery with it, etc. Because most blind pupils were new to working with a mouse device, we explained the tactile exploration analogy. Tactile maps can be explored by feeling lines and shapes with two or more fingers and the mouse works quite the same. The position of the mouse is similar to the position of the fingers when exploring tactile imagery.

While playing with the VTPlayer, the blind pupils were introduced to the software that is included in the VTPlayer package. Using this software for about 30 minutes, the blind pupils learned the basic use of the mouse. The following programs were explored: BullsEye, Hide and Seek, Duck Shoot, and Tactile Maps Series: Europe.[14] The latter is the most mentally challenging program and demanded a combination of motor and cognitive proficiencies. The VTPlayer RE*quest* focused on learning geography with the VTPlayer.

Because obviously these pupils could not read the questionnaire, we presented the items in digital audio format (in Dutch). Each item was separately recorded and stored as a small audio file. We wrote some software that selected and played items in pseudo-random order, different for each individual pupil (Van den Berg & Hoorn, 2005, Tech. Rep. [CD]). To avoid fatiguing effects, the audio VTPlayer RE*quest* was divided into four blocks, followed by a small pause. The blind pupils heard the items over their headphones and submitted a score by pressing one of the keys 1 to 6 on a standard keyboard, simulating a conventional 6-point rating scale. The zero was not used because in the standard keyboard layout this key lies out of range. The pupils automatically received acoustic feedback on their response by repeating the key they pressed. Scores had to be confirmed by pressing Enter in order to proceed. Data were automatically stored. Illegal or out-of-range responses were replied to with warning feedback and could be replaced by the proper response. Running the audio VTPlayer RE*quest* was self-paced and took about 30 minutes. A transcription of the questionnaire is available in Appendix 4.3.

[14] http://www.virtouch2.com/Products.htm

Sample requirements in the VTPlayer case

- Must (Personal)
  - Straight lines
  - Simple forms
  - Presentation of the most important features
  - Navigation with sound effects
  - Different size hatches
  - Spoken introduction to image
  - Clickable features
  - Large size images
- Won’t (Personal)
  - Curved lines
  - Complicated forms
  - Irrelevant details
  - Navigation without sound effects
  - All lines equal width
  - Same hatches used more often
  - Lines within images
  - Small images

Sample goals in the VTPlayer case

- Approach (Personal)
  - Easy to follow
  - Better understanding
  - Learning support
  - Gaining insight
  - Know what is important
  - Know what information belongs where
  - Independent learning
- Avoid (Personal)
  - Difficult to follow
  - Loss of concentration
  - Confusion
  - Being at wrong position in screen layout
  - Information overload
  - Lost in navigation
  - Different forms looking the same

4.6.1.4 Measurements

Scale construction in the VTPlayer case

The VTPlayer RE*quest* presented 4 items about the relevance of the VTPlayer to personal study goals. Examples are "The VTPlayer is important for my concentration" and "The VTPlayer is influential for my grades."

The Stakeholders' Needs scale consisted of 48 items. These items linked personal goals to graphical and information requirements. The items were presented as requirements statements about the VTPlayer and related geography software. An example item is "Information by speech improves my independence." "Information by speech" was a Must requirement, "improves" induced positive valence (Support), and "independence" was a personal goal to achieve (Approach). Items were scored for agreement on 6-point rating scales (1= completely disagree, 6= completely agree).

4.6.2 Analysis and Results

Data were entered in an SPSS 12.0 data matrix for statistical analysis. For a detailed analysis, see (Van den Berg & Hoorn, 2005, Tech. Rep. [CD]).

4.6.2.1 Scale Analysis

In doing scale analysis, I followed the same procedures and criteria as described in Section 4.3.2.1. To achieve acceptable reliability I deleted one item (on mood) from the Relevance scale (Standardized Cronbach's alpha= .65). The VTPlayer was most important to gain insight (M= 4.20, SD= 1.32), followed by learning pleasure (M= 3.47, SD= 1.06), and grades (M= 2.13, SD= 1.13).

Standardized Cronbach's alpha of the revised sub scales of the Stakeholders' Needs scale was between .70 and .83, which is good. Scale length was between 2 and 4 items. In all, measurement quality in this study was improved compared to some of the previous studies. For more details, see (Van den Berg & Hoorn, 2005, Tech. Rep. [CD]).

4.6.2.2 Agreement to VTPlayer Requirements and Goals

A 2*3 MANOVA of 2 (Item Type: Indicative vs. Contra-indicative) (within-subjects) * 3 (Scales: Requirements vs. Valence vs. Personal Goals) was run on the grand mean level of agreement to requirements statements (Figure 10). For the complete analysis, see (Van den Berg & Hoorn, 2005, Tech. Rep. [CD]).

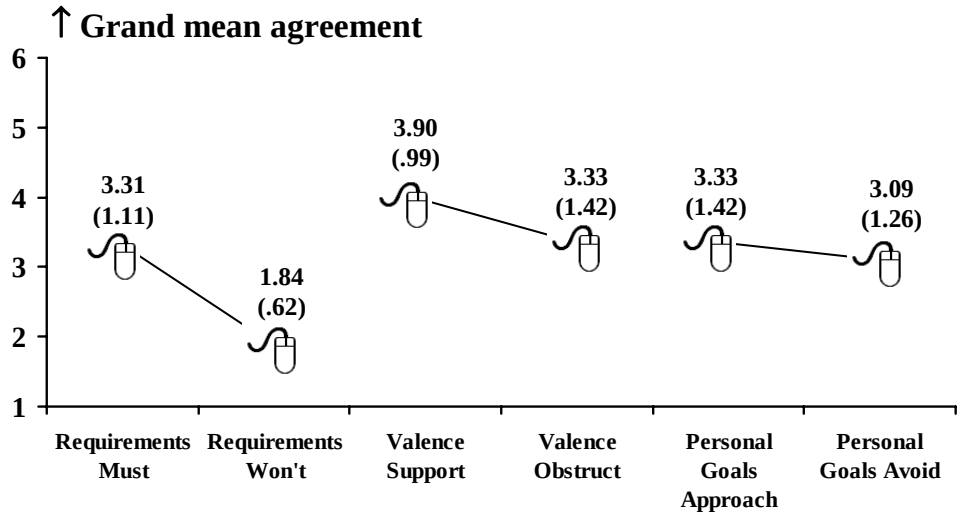


Figure 10: Grand mean average agreement to the 6 sub scales of Stakeholders' Needs (N= 15). Standard deviations are in parentheses.

The interaction between Item Type and Scales was significant, Pillai's Trace= .56, $F_{(2,13)}$= 8.21, p= .005. This effect was rooted in the difference between Indicative and Contra-indicative items of Requirements vs. Personal Goals, parameter coefficient= 1.22, t= 3.45, p= .004, η_p^2= .46. The difference in agreement between Must (M_{Must}= 3.31, SD= 1.11) and Won't ($M_{Won't}$= 1.84, SD= .62) requirements was larger than the difference between Personal Goals Approach ($M_{Approach}$= 3.33, SD= 1.42) and Personal Goals Avoid (M_{Avoid}= 3.09, SD= 1.26). Won't requirements yielded the lowest levels of agreement.

In addition, the main effect of Scales was also significant, Pillai's Trace= .67, $F_{(2,13)}$= 13.06, p= .001. In particular, the level of agreement to Requirements (M_{Reqs}= 2.58) was much lower than to Valence ($M_{Valence}$= 3.62), parameter coefficient= -2.08, t= -5.24, p< .001, η_p^2= .66. As a trend, the level of agreement to Valence ($M_{Valence}$= 3.62) was higher than to Personal Goals (M_{Goals}= 3.21), parameter coefficient= .81, t= 2.15, p< .049. However, this difference was not significant according to Bonferroni (α= .05/3 ≈ .017).

Adding Sex as a control factor raised insignificant effects. Yet, there was an effect of Age (13-17) as a covariate of Valence Support, $F_{(1,14)}$= 10.01, p= .007, indicating that the higher the age, the less support was expected (r= -.66), which is understandable.

In conclusion, the sub scales were capable of raising strong significant effects, particularly when Requirements and Personal Goals were involved.

4.6.2.3 Discussion of the Effects on Agreement in the VTPlayer Case

The significant interaction between Item Type and Scales showed that the scores went into the expected direction. The blind pupils indicated higher levels of agreement to Must than Won't, to Support than Obstruct, and to

Approach than Avoid. Valence was mainly active in combination with Requirements.

4.6.2.4 VTPlayer: Personal-viewpoint Requirements in Relation to Personal Goals Revisited

In Study 4, I tested the predictions P1-P8, following the same procedures as in Study 3 (Section 4.5.2.4).

Explaining Requirements Must

Regression (method Enter) of Personal Goals Avoid on Requirements Must revealed that Personal Goals Avoid indeed contributed significantly to the Requirements Must variability, R^2= .29, R^2_{adj}= .23, $F_{(1,13)}$= 5.18, p= .040, standardized β= -.53.

As a control, multiple regression was performed (method Enter) of Support and Obstruct (entered in the first step) and Personal Goals Avoid and Personal Goals Approach (in the second step) on the Requirements Must measure.

Support and Obstruct could significantly explain Requirements Must, R^2= .44, R^2_{adj}= .35, $F_{(2,12)}$= 4.78, p= .030. This was mainly due to Obstruct, standardized β= .66, t= 3.09, p= .009, $r_{partial}$= .67, r_{part}= .67. Support could not significantly surpass this contribution. Additionally, Personal Goals Avoid was not strong enough ($F_{change(1,11)}$= 4.02, p= .07) to contribute above and beyond Obstruct.

To test whether Obstruct could be considered a mediator between Must and Avoid, I ran another regression analysis of Personal Goals Avoid on Obstruct. However, no significant contribution was found, $F_{(1,13)}$= 1.03, p= .329. Thus, Valence Obstruct seemed to be an independent predictor of the Requirements Must variability.

Explaining Requirements Won't

I performed regression analysis (method Enter) of Personal Goals Approach on Requirements Won't and found that Personal Goals Approach contributed significantly to Requirements Won't, R^2= .36, R^2_{adj}= .31, $F_{(1,13)}$= 7.30, p= .018, standardized β= -.60.

As a control, I performed multiple regression analysis (Enter) on Requirements Won't with Valence Support and Obstruct entered in the first step and Personal Goals Avoid and Personal Goals Approach in the second step.

None of the predictors were strong enough to make a significant contribution. For Support and Obstruct, $F_{(2,12)}$= 3.44, p > .05. The additional effect of Approach and Avoid was also insignificant, $F_{change} < 1$.

In other words, Personal Goals Approach considerably explained Requirements Won't but was not so strong that its contribution went above and beyond Support and Obstruct. However, Support and Obstruct were themselves not capable of explaining Won't requirements either. This excludes the possibility that Support and Obstruct mediated the relation between Personal

Goals Approach and Requirements Won't. Nevertheless, the MANOVA showed that Valence did yield significant effects. Thus, Support should be regarded as a moderator of the relation between Personal Goals Approach and Requirements Won't.

4.6.2.5 Discussion of the Relation between Personal-viewpoint Requirements and Personal Goals Revisited

I confronted the results of Study 4 on the blind pupils evaluating the VTPlayer with the predictions P1 to P8.

(P1) (Approach→Won't) was confirmed. Personal Goals Approach independently predicted Requirements Won't (R^2_{adj}= .31, standardized β= -.60).

(P2) (Avoid→Must) was confirmed. Personal Goals Avoid independently predicted Requirements Must (R^2_{adj}= .23, standardized β= -.53).

(P3) ¬(Avoid→Won't) was confirmed. Personal Goals Avoid offered hardly any additional predictive power to the Won't variability over Personal Goals Approach.

(P4) ¬(Approach→Must) was confirmed. Personal Goals Approach provided little or no additional predictive power to the Must variability compared to the contribution of Personal Goals Avoid.

(P5) (Approach$\overset{\text{Support}}{\not\downarrow}$Won't) was confirmed. Valence Support did not have significant explanatory power for the Requirements Won't variability. Yet, MANOVA did reveal significant effects of Valence on the level of agreement.

(P6) (Avoid$\overset{\text{Obstruct}}{\not\downarrow}$Must) was rejected (see P8).

(P7) (Approach→Support→Won't) was rejected (see P5).

(P8) (Avoid→Obstruct→Must) was rejected. Although Valence Obstruct predicted (standardized β= .66, $r_{partial}$= .67, r_{part}= .67) the Must variability while annihilating the predictive power of Personal Goals Avoid, in its turn, Personal Goals Avoid did not significantly explain Valence Obstruct.

As in the COTS PCs case (Study 3), all the predictions were confirmed or rejected by the VTPlayer data according to the goals-to-requirements chiasm (Figure 11), except for, again, P6.

This time, however, the rejection was more severe because Valence Obstruct now served as an independent predictor of Must Requirements. And as stated in Section 4, all other relations than predicted by P1 to P8 count as model refutation.

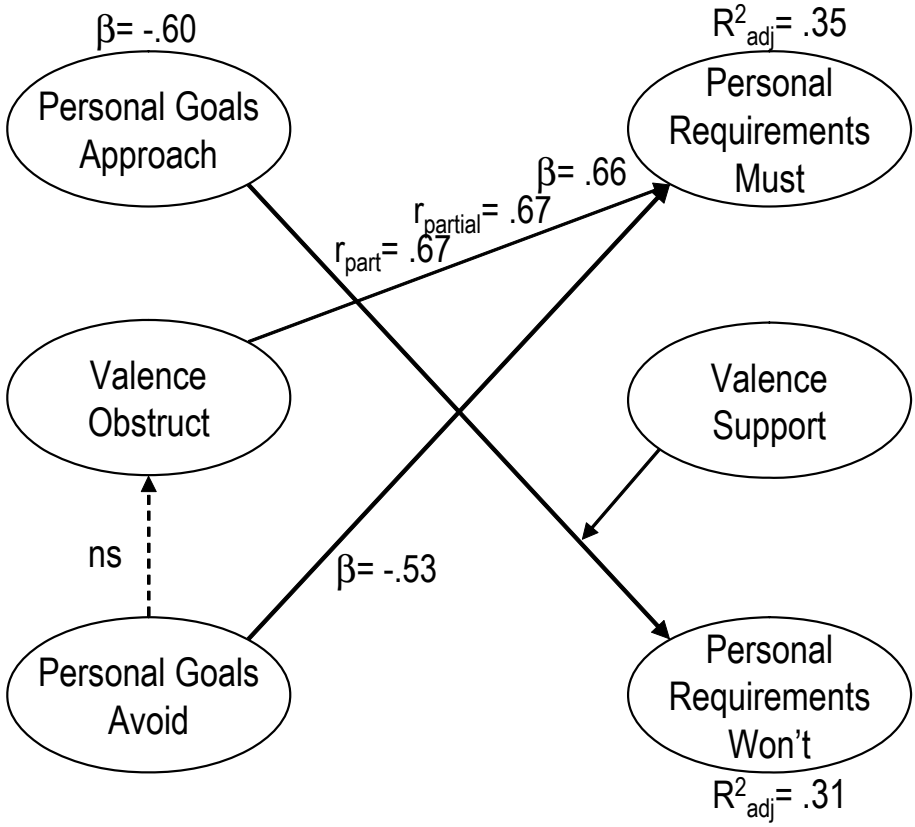


Figure 11: Goals to Approach explain variability in the Won't Requirements, Goals to Avoid in the Must Requirements. Valence Obstruct served as predictor of Must variability.[15]

4.6.3 VTPlayer: Practitioner's Perspective

For the Bartimeus Foundation and VISIO it is important to have academic evaluations of their learning methods and devices. The market is tight and they need independent advice on the usefulness of products. For the same reasons, developers such as VirTouch want to stay in touch with science to have a pre-competitive advantage over other suppliers. They use the scientific results to improve their products and tailor the design to their customers.

Thus far, the practitioners working with the different versions of the RE-*quest* worked with adult stakeholders who were sighted. Writing the VTPlayer RE*quest* for blind adolescents was something quite different, so Jelle van den Berg experienced. During the interviews and participatory observations with several groups of 15 year olds, this practitioner noticed that:

"Their skill in independently interacting with PCs differs but except for a few, they all are used to working with computers. It is a good thing to do such observations: One of the things that came out of it is that the coaching is quite important. Sometimes, while exploring an Office application that is mal-

[15] The Beta weights and correlations of the two sub models ((Avoid→Must, Obstruct→Must) and Approach→Won't) should not be compared.

adapted to these blind pupils, they encounter problems that they cannot solve by themselves. In such instances, their problem-solving capacities fall short to independently figure things out. However, if only they get a small cue into the right direction, things go much better so that's something I need to pay attention to in the RE*quest* as well as in the prototype. Something else that catches the eye is the long learning curve of the pupils: Learning a new program takes a lot of time but when you take this time, working with the program goes brilliantly" (personal e-mail communication, Sept. 20, 2004).

Another point was to get enough stakeholders to participate in the field experiment. The practitioner wrote: "I am also connecting Sensis, the third organization for the visual impaired in The Netherlands. I hope for more pupils. The main point actually is that [the government] tries to send most blind children to the regular schools, supervised by someone who monitors the integration of the blind. Thus, there are enough test persons around, but these do not come to an organization such as Bartimeus or Visio" (personal e-mail communication, Oct. 20, 2004).

Building the VTPlayer RE*quest* was not an easy thing to do. On the technicalities of constructing the Stakeholders' Needs scale, the practitioner commented:

"Perhaps nice for you as the builder of the test to know the approach I took. First, I gathered a lot of requirements that I drew from the scientific literature, partly derived from the current system's features, partly from the interviews, and partly from the conclusions that I drew from my own sources. Here I began to separate between information supply and the structure of the images. Then I specified the goals but it was quite hard – in view of the topic – to think up negative goals. After a lot of overhaul, I started to make combinations [of goals and requirements] and all possible combinations I jotted down. Next, I put priorities to the combinations. It turned out that I actually had two goals per requirement that were important. But I already had too many requirements so a 1 : 2 ratio was not possible without exploding the number of items. I then decided to take the combination with the highest priority. This was the first basis for the PRE*quest*, in which many goals were used twice. The big puzzle really began only then and I had great difficulty in using each goal but once and keeping the questionnaire logical and balanced as a whole. I often found that I already had enough items of one category (e.g., Must-Obstruct-Avoid), after which I had to look whether I could make a negative variant of a goal, found it, but did not like it because that specific requirement needed a positive goal or that the negative variant did not cover the meaning of the original goal in the same way" (personal e-mail communication, Feb. 2, 2005).

4.7 General Conclusions and Discussion

Requirements change as the situation in which information systems function evolves (Alves & Finkelstein, 2002). Situations change as a result of certain

events, a change of tasks, adopting another business or personal model or a change in (organizational) culture (Damian et al., 2004). Stakeholders call for or dismiss requirements and errors should be repaired (Alves & Finkelstein, 2002). However, different stakeholders may have conflicting requirements (Spanoudakis, Finkelstein, & Till, 1999), which points at opposing goals or different means of achieving them in the new situation. While situations, and subsequently, requirements develop, uncertainty can be managed and the new situation controlled as soon as requirements are again agreed-upon (Alves & Finkelstein, 2002). To manage a change requirement, goals are fundamental for discovering conflicts among (the new) requirements (Van Lamsweerde, 2004). "Goals provide the rationale for requirements i.e. requirements represent one particular way to achieve high-level goals" (Alves & Finkelstein, 2002) (e.g., strategic business goals).

4.7.1 The Goals-to-Requirements Chiasm

My primary aim to constitute an account of requirements change was accomplished in a way I did not foresee beforehand. I started from the common assumption that the new things stakeholders wanted on a system were demanded by changes in goals stakeholders wanted to achieve. As a counterpart, I believed that what should go off a system was directed by situations stakeholders wanted to avoid. I added the concept of valence to relate changes in agreement to requirements to positive or negative expectations of using the system in the future.

Contrary to accepted belief, however, I found quite a reverse relationship, which I coined the goals-to-requirements chiasm (χ-effect), which happened when requirements and goals were evaluated from the same stakeholder viewpoint (Chapter 3). It turned out that for the most part, goals to approach explained the variability in agreement to won't requirements, whereas goals to avoid explained the variability in agreement to must requirements. That is, goals that stakeholders desire dictate changes in the won't requirements whereas things they fear dictate changes in the must requirements. In all studies except one the straight relationships between approach and must as well as avoid and won't were insignificant.

What happened was probably this. There are requirements a system must have. A car, for instance, should have tires, a steering wheel, and breaks. There is not much discussion about these must haves. Every next version of a car has the same features because the related goals are positively charged (to advance smoothly, to change direction, and to stop, respectively). These positive must requirements remain stable throughout changing situations because they are connected to positive goals. Thus, variability in agreement to such requirements is little. There is a ceiling effect of agreement. Everybody wants this type of must requirement throughout changing situations.

There are also requirements nobody wants on a future system for reasons that remain stable. The car should not have a steam engine (too slow), solid

tires (not comfy), or a steering rod (unhandy). In Study 3, nobody wanted green-on-black cathode ray tubes for a monitor because it was ugly and bad for the eyes. Here we have a bottom effect of disagreement. Everybody disagrees on such 'negative' won't requirements because these are connected to goals to avoid (negative goals). Throughout changing situations, the variability in disagreement to these won't requirements remains little. There is hardly any variance to explain.

However, the volatile requirements, *those requirements most susceptible to change are the ones that are connected to goals of opposite polarity.* There may be quite some dispute about installing virus-protection software in printers that do not communicate with the outside world. For the managers of the New York business company, this was a must because they feared the damage that malicious code can inflict upon a system. Yet, other stakeholders such as the supplier (Océ Technologies) would argue that it is ridiculous to do so because those printers cannot be infected. The 'positive' requirement ("We agree to virus-protection") is connected to and simultaneously clashes with a negative goal state ("We disagree to possible damage"). Because of these adverse effects, variability in agreement is large and consequently, change requirements are triggered easily.

Of course, the reverse structure works in the same way. Not migrating e-Synergy to Linux (the won't requirement) may preserve the European market (a desired goal) but leaves the door closed to China and other Eastern markets (Van Nieuwland, personal communication, Nov. 17, 2004). Moreover, Linux is the more reliable shell. Therefore, Exact is contemplating to migrate e-Synergy from Microsoft to Linux some day in the near future (ibid.). Again, variability in agreement to the 'negative' requirement is high ("I disagree to Linux") because it is connected and clashes with a positive goal ("I agree to profit maximization"). Thus, the requirement is sensitive to change.

Taken in unison, requirements that must be on a system have a baseline agreement that is pushed down by the disagreement of the stakeholder to an undesired future situation. Mirroring this, requirements of things the system won't have, evoke a baseline disagreement that is pulled up by the agreement of the stakeholder to a desired future situation. Ergo, requirements change.

The χ-effect occurred throughout an array of different conditions. If the viewpoints on goals and requirements are the same, it occurs for personal as well as business goals. It occurs for different types of goals (e.g., efficiency, budget, or learning effect), for different types of systems and requirements (e.g., capacity and warehouse management, COTS PCs, and a tactile mouse). It happens for different types of stakeholders (e.g., police officers, provincial managers, CHI experts, or blind college students). Stakeholders may be from different countries (in Study 3, Holland, Poland, India, Israel, Romania, and Russia). It does not matter whether stakeholders are novices (Study 4) or experts (Study 3). The χ-effect occurred throughout different modalities (written or spoken questionnaires). Unreliability of measures is hardly a problem

(Study 2) (which is by no means an argument for sloppy measurement). The χ-effect is independent of sex, function, or other background variables. Only in Study 4, the system was deemed less helpful when the users were older. Yet, the really critical thing was that viewpoints on requirements and goals should be aligned.

4.7.2 The Requirements-analysis Rift

In a series of four studies, I was capable of replicating the χ-effect. The Personal View data set in Study 1, however, failed to show the crossed relations. In hindsight, this can be explained from the viewpoint-alignment line of reasoning I called the requirements-analysis rift (Chapter 3). In Study 1, there was a mismatch of views deliberately induced to the stakeholders. In one condition, I wanted to see in how far the future users from a personal point of view would evaluate management-viewpoint requirements. As can be seen in Figure 1, from a personal viewpoint everything was upside down. Personally, the officers thought that the management's requirements were not indisputable. There was a lot of disagreement to what the managers proposed as must requirements. The officers thought they were won't haves. Conversely, what the managers proposed as won't haves was seen as must haves by the officers. The same happened for the personal goals to avoid or achieve with the system (Figure 2). In this situation, no ceiling or bottom effects in the variability of agreement to the requirements statements occurred because everything was under debate. Probably, in situations of ongoing negotiation on the positive or negative status of requirements and goals, change requirements arrive from the straight relationships as depicted in Figure 4. The χ-effect evaporates when viewpoints clash, that is, when the requirements-analysis rift emerges.

4.7.3 The Status of Valence

Juliet is like the sun and so is valence, in particular, when it comes to positive expectations about omitting certain requirements from the system to accomplish a desired goal ("Get rid of Windows for a stable system"). In all five data sets, Valence Support acted as a moderator and in four data sets between Goals to Approach and Won't Requirements. This 'leg' of the χ-effect was unproblematic in that the relation between Approach and Won't could always be established while Support regulated the agreement to the requirements statements but as a factor from the outside. Like the sun can enhance but is not necessarily the reason for being happy, Valence Support influenced the level of agreement.

Valence Obstruct is like a wildcard. It is hard to determine what value it will take. In four cases, it served as a moderator (Figures 4, 7, 9, 11). In one case it was a mediator between Must Requirements and Goals Avoid (Figure

9). In two cases, Valence Obstruct was an independent predictor of the Must variability (Figure 5 and 11), once in connection to Approach (Figure 5).

Whatever role Valence Obstruct took, it was in four cases directed at the Must Requirements (Figures 5, 7, 9, 11). Because three times (Figures 7, 9, 11) I could establish the second 'leg' of the χ-effect between Avoid and Must, I assume that Valence Obstruct is active for this relationship. In fact, it was never active for the relation between Approach and Won't. Thus, negative expectations are influencing or sometimes even explaining variability in agreement to requirements that are wanted to ward off something frightening (e.g., a virus infection). It seems that in a way, stakeholders do not trust too much the solutions that are offered to them to deter danger. Figure 12 depicts the χ-effect as I generally conceive of it.

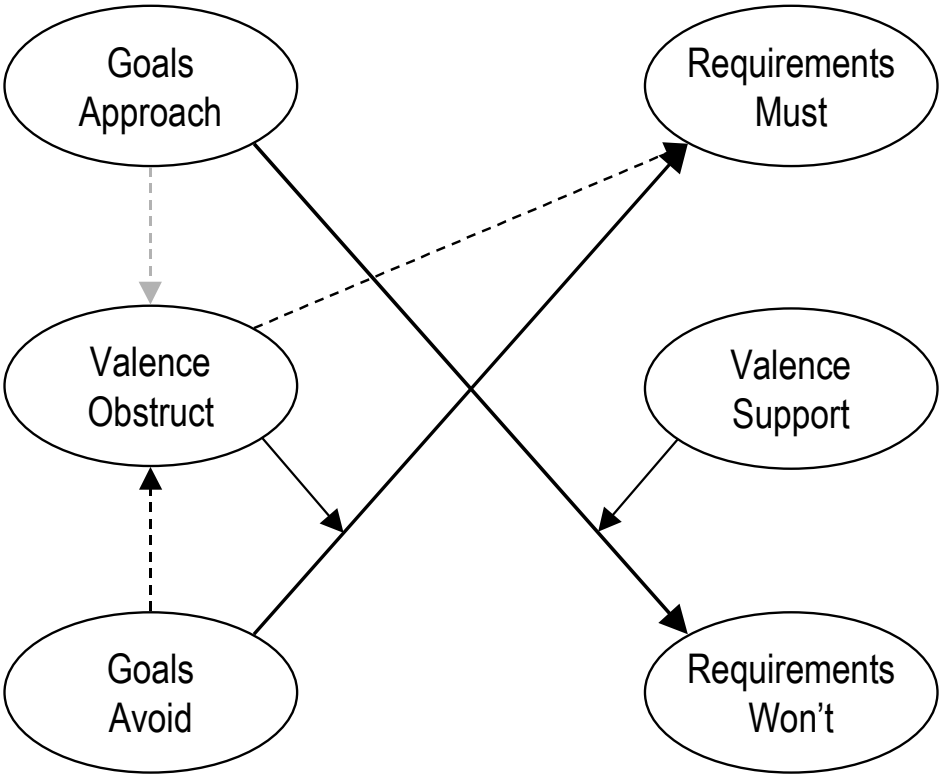


Figure 12: The χ-effect. Valence Obstruct is a moderator, sometimes an independent predictor, and in one case a mediator between Avoid and Must.

4.7.4 Methodological Challenges

I could take pride in that all my findings were significant while using very small samples of stakeholders (14 to 18 people). However, the problem about the status of Valence Obstruct, the missing relation between Avoid and Must

in Study 1 (Figure 5), or the sometimes-poor measurement quality (Study 2) may all be due to that very point.

It used to be a merit that with smaller samples, the number of items was reduced – in the cases I presented; only 2 to 4 items survived the scale analyses. Yet, certain authors state that this assumption may be improper (e.g., Guadagnoli & Velicer, 1988; Marsh et al., 1998; Velicer & Fava, 1998; MacCallum, 1999). Sure, variables are less stable with fewer observations so that responses to more items by more stakeholders will reduce variability.

However, Cohen and Cohen (1975) already showed that the stronger the effect, the fewer observations are needed. In my studies, then, the problem is that I did not know the strength of the effects beforehand. In the CMS case, for example, the relation between Approach and Won't (Figure 5) in the Business View was a meager .16. According to (Cohen & Cohen, 1975), I should have used 75 stakeholders to be 95% (at $\alpha = .05$) that the relation was at least .40. Obviously, I did not. Yet, this was the first time I established one leg of the χ-effect, moreover, in the somewhat weird situation that requirements and goals were from the management whereas the judges were plain employees. Therefore, I felt the urge to replicate the findings, for example, in the LWMS case. If we look at the strength of the relationships in the LWMS case (.90 and .70), small samples seem to be warranted. Cohen and Cohen (1975) teach that if we want to be 95% sure (at $\alpha = .05$) that the strength of the relationship is at least .70, we need 19 stakeholders (we came one short). Thus, yes, larger sample sizes are preferable but I compensated for that by replicating the relationships. In addition and in hindsight, we seem to deal with pretty robust effects because these occurred independently of system, setting, or socio-demographic variables, reducing the severity of the small samples problem. In fact, from a whole different angle and with a different type of questionnaire, I unintentionally established the χ-effect yet another time for operation-room machines in the hospital (Chapter 5).

The challenge, then, is to find a large enough company in which one group of stakeholders can be used to test the questionnaire items for psychometric quality. The other group should consist of a few hundred people to test the predictions P1 to P8. This would also solve the problem that I actually tested two sub models (the two 'legs') within the χ-effect. This makes it hard to compare the relative explanatory power of the predictors in the one sub model with that in the other sub model. A large enough sample size allows for Structural Equation techniques, which account for more sources of variance and can estimate the fit of the model as a whole.

4.7.5 Message to and Testimonies from Practitioners

Do you want to anticipate or merely react to change? The most important information an IT practitioner could extract from a system's stakeholders are covered by four questions. What are the things in life or work that you do not

want? What can the system offer to avoid those things? What are the things in life or work that you do want? What should the system *not* have in order to support that? In view of the relative importance of features the future system should not have, it seems that analysis of the won't requirements is underestimated in industrial practice.

Stakeholders maintain a baseline agreement to must requirements, which is regulated by the disagreement to the 'threat' to goals in the future. According to Jo Geraedts, Industrial Design Dept., Océ-Technologies (personal communication, Nov. 11, 2004), this is the "cover your ass" attitude many stakeholders take when they are asked to outline the future state of the system or the work environment. In opposition, won't requirements evoke a baseline disagreement that is governed by agreement to possible support of desirable goals in the future. According to Geraedts, this is the "make life easier" attitude that stakeholders also wish to maintain. The practical substrate of the formal χ-effect, then, is that stakeholders work from an attitudinal disposition that says "make life easier, while covering your ass."

In this chapter, I added practitioner's perspectives on using an elaborative method to validate the observed agreement to requirements. From these experiences, the case for the industrial audience is probably simple: The proposed methods are too dense, too difficult to understand, and too labor intensive to get used frequently. But then again, they don't need to. This problem of the *business utility of scientific methods* (Chapter 2) is based on a false impression. It is not the task of businesses to do such research; that is what the academy is for. It is the academic task to provide solid research results, which can later on be translated into focal points of RE in practice (e.g., the first paragraph of this section), possibly apprehended with more lightweight approaches.

Related to this utility problem is that well-established methods in the RE literature such as VORD (Viewpoint Oriented Requirements Definition) (Kotonya & Sommerville, 1992), KAOS GRAIL (Bertrand et al., 1998), ScenIC (Potts, 1999), and in industrial practice Volere (Robertson & Robertson, 2004) could all do the job just as well as the statistically intensive techniques employed here. How to justify the extra 'bangs for bucks' that a statistically motivated approach delivers (Chapter 2)? Well, it is not a matter of producing elegant statistics to outdo existing RE methods. The good news is exactly that Volere and other methods *can* establish the same results. Point is, however, that they never did and that the statistical approach advanced in this paper uncovered relationships and possibilities never contemplated before. As argued in Chapter 2, it is the academic task to develop a reliable, valid, and general theory of requirements (here, a contribution to the theory of requirements change), which may guide the follow-up in the industry by means of statistically less intensive approaches such as Volere or VORD.

Simon van Dam, Vice President of VirTouch Ltd., Israel, is one of the developers of the VTPlayer Braille mouse. He found our group on the Web and

interested in our approach, he came to Amsterdam to discuss our work and hear our ideas about improving the VTPlayer before making us VirTouch's official research partner (personal communication, Jan. 21, 2004).

Guido Fambach is Director Education of The Mediator Group, a Dutch firm that develops the Didactor e-learning environment.[16] As a member of the industrial supervising committee of this here project of the Ministry of Economic Affairs, he affirmed that he intended to use our insights for his practical work and to pose the questions as put in the first paragraph of this section (personal communication, Nov. 17, 2004).

SIGCHI.NL (Oct. 13, 2005) is a meeting where business and science exchange information and experiences. While I presented some of the results (Hoorn, 2005b) discussed in the present chapter, I wanted an informal count whether business people and IT practitioners acknowledged the χ-effect as something to reckon with in RE. When the first author said: "Everybody raise hands because otherwise I don't get published," nobody moved a finger. The whole room was discussing whether it would help ("Probably not") if I asked them to do something positive to keep me from something negative.

References

Alves, C., & Finkelstein, A. (2002). Challenges in COTS-decision making: A goal-driven requirements engineering perspective. *Workshop on Software Engineering Decision Support* (Ischia IT, 2002).

Alves, C., & Finkelstein, A. (2003). Investigating conflicts in COTS decision-making. *International Journal of Software Engineering and Knowledge Engineering, 13*, 1-21.

Anton, A. I., Cracken, W., & Potts, C. (1994). Goal decomposition and scenario analysis in business process reengineering, in *Proceedings of CAiSE '94* (Utrecht NL, 1994).

Anton, A. I., Earp, J. B., Potts, C., & Alspaugh, T. A. (2001). The role of policy and stakeholder privacy values in requirements engineering, in *Proceedings of RE'01: International Symposium on Requirements Engineering* (Toronto, Canada, 2001).

Baron, R., & Kenny, D. A. (1986). The moderator-mediator variabele distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*, 6, 1173-1182.

Bertrand, P., Darimont, R., Delor, E., Massonet, P., & Van Lamsweerde, A. (1998). GRAIL/KAOS: An environment for goal driven requirements engineering. *Proceedings of ICSE'98 - 20th International Conference on Software Engineering*. Kyoto, Japan: IEEE-ACM.

Cohen, J. & Cohen, P. (1975). *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences.* Erlbaum, Hillsdale NJ.

Damian, D., Zowghi, D., Vaidyanathasamy, L., and Pal, Y. (2004). An industrial case study of immediate benefits of requirements engineering process improvement at the Australian Center for Unisys Software. *Empirical Software Engineering, 9*, 1-2, 45-75.

[16] www.mediatorgroup.com/02_elearning.php

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly 13*, 3, 319–339.

Dillman, D. A. *Mail and Internet Surveys: The Tailored Design Method.* Wiley, New York NY, 1999.

eRA (2002). *Program Portal Scope Document v1.4.* (Tech. Rep.). National Institutes of Health, Bethesda MD, 2002. Available at http://era.nih.gov/Docs/Pgm_Por_Scope_4-02.doc

Frijda, N. H. (1986). *The Emotions.* Cambridge University Press, New York NY, 1986.

Guadagnoli, E., & Velicer, W. F. (1988), Stability of component patterns: A simulation study. *Psychological Bulletin, 103* (1988), 265-275.

Guttman, L. An outline of some new methodology for social research. *Public Opinion Quarterly, 18* (1954), 395-404.

Guttman, L. A faceted definition of intelligence. *Studies in Psychology, 14* (1965), 166-181.

Hoorn, J. F. (2004). *Validating Stakeholders' Needs of a Logistic Warehouse Management System.* (Tech. Rep.) [CD]. Amsterdam: Free University.

Hoorn, J. F. (2005a). *Computers Off-the-Shelf: The CHI REquest.* (Tech. Rep.) [CD]. Amsterdam: Free University.

Hoorn, J. F. (2005b). Sources of requirements change. A goal and viewpoints-driven explanation. *SIGCHI.NL* The Hague, Netherlands, October 13, 2005. New York: ACM Press.

Hoorn, J. F., & Kok, E. (2005). *Requirements Validation of a Capacity Management System at the Dutch Police Force*. (Tech. Rep.) [CD]. Amsterdam: Free University.

Hoorn, J. F., Konijn, E. A., Van Vliet, H., & Van der Veer, G. (2005). Goal-oriented RE for handling change requirements: An explanation of what stakeholders try to avoid and what they try to achieve. In K. Cox, E. Dubois, Y. Pigneur, J. Verner, S. Bleistein, A.M. Davis, & R. Wieringa (Eds.), *Proceedings of the International Workshop on Requirements Engineering for Business Need and IT Alignment (REBNITA'05) in Conjunction with the Thirteenth IEEE Requirements Engineering Conference (RE'05)* (Paris, France, 2005).

Hoorn, J. F., Konijn, E. A., Van Vliet, H., & Van der Veer, G. (2006). Requirements change: Fears dictate the must haves; desires the won't haves. *Journal of Systems and Software*. In press.

Jordan, B. (1996). Ethnographic workplace studies and computer supported cooperative work, in D. Shapiro et al. (Eds.) *The Design of Computer-supported Cooperative Work and Groupware Systems.* North Holland/Elsevier, Amsterdam NL (1996), 17-42.

Kotonya, G., & Sommerville, I. (1992).Viewpoints for requirements definition. *Software Engineering Journal, 7*, 6, 375-387.

Lehman, M. M. (1996). Laws of software evolution revisited. Lecture notes in Computer Science, 1149. *Proceedings of the 5th European Workshop on Software Process Technology* (pp. 108-124). Springer-Verlag: London UK.

Letier, E., & Van Lamsweerde, A. (2001). Agent-based tactics for goal-oriented requirements elaboration, in *Proceedings of 24th International Conference on Software Engineering* (Orlando FL, 2001).

MacCallum, R. C., Widaman, K.F., Zhang, S. and Hong, S. (1999). Sample size in factor analysis. *Psychological Methods, 4*, 1, 84-99.

Marsh, H. W., Hau, K. T., Balla, J. R., and Grayson, D. (1998). Is more ever too much?: The number of indicators per factor in Confirmatory Factor Analysis. *Multivariate Behavioral Research, 33*, 2 (1998), 181-220.

Norman, D. (1998). *The Invisible Computer*. MIT Press: Cambridge MI.

Ohlsson, M. C., & Runeson, P. (2002). Experience from replicating empirical studies on prediction models. *Proceedings of the 8th IEEE International Symposium on Software Metrics (METRICS'02)*, Ottawa, Canada (pp. 217-226). Los Alamitos, CA: IEEE Computer Society.

Potts, C. (1995). Using schematic scenarios to understand user needs, in *Proceedings of DIS'95 - ACM Symposium on Designing Interactive Systems: Processes, Practices and Techniques* (Michigan MI).

Potts, C. (1999). ScenIC: A strategy for inquiry-driven requirements determination. *Proceedings of RE'99: International Symposium on Requirements Engineering.* Limerick, Ireland.

Price J., & Cybulski J. (2004). Influence of stakeholder communication on consensus making in requirements negotiation, in *Proceedings of AWRE* (Adelaide, Australia).

Pronk, S. (2004). *From Static Store to Organized Warehouse.* (Tech. Rep.) [CD]. Amsterdam: Free University.

Robertson, J., & Robertson, S. (2004). VOLERE: Requirements specification template. Edition 10.1. Atlantic Systems Guild.

Robinson, W. N., and Pawlowski, S. (1999). Managing requirements inconsistency with development goal monitors. *IEEE Transactions on Software Engineering,* (Nov/Dec 1999).

Sommerville, I., & Sawyer, P. (1997). Viewpoints: Principles, Problems and a Practical Approach to Requirements Engineering. *Annals of Software Engineering, 3*, 101-130.

Spanoudakis, G., Finkelstein, A., & Till, D. (1999). Overlaps in requirements engineering. *Automated Software Engineering, 6*, 2, 171-198.

Sutcliffe, A. (2002). *User-centred Requirements Engineering. Theory and Practice.* Springer, London UK.

Van den Berg, J. Y., & Hoorn, J. F. (2005). *VTPlayer REquest.* (Tech. Rep.) [CD]. Amsterdam: Free University.

Van Lamsweerde, A. (2004). Goal-Oriented Requirements Engineering: A Roundtrip from Research to Practice, in *Proceedings of the Requirements Engineering Conference, 12th IEEE International (RE'04)* (Kyoto JP, 2004), Volume 00, 4-7.

Van Lamsweerde, A., & Letier, E. (2000). Handling obstacles in goal-oriented requirements engineering. *Software Engineering, 26*, 10, 978-1005.

Velicer, W. F., and Fava, J. L. (1998). Effects of variable and subject sampling on factor pattern recovery. *Psychological Methods, 3*, 2, 231-251.

Appendix 4.1

LWMS Re*quest*[17]

SYSTEEMVEREISTEN

De volgende uitspraken betreffen een aantal eigenschappen die het nieuwe magazijn-beheersysteem volgens u zou moeten hebben. Geef alstublieft aan in hoeverre u het daarmee eens bent door een getal te omcirkelen 0 (helemaal oneens) - 5 (helemaal eens).

Hier volgt een aantal uitspraken over de bestelprocedure in de mogelijk toekomstige situatie.
De medewerker…

Qarc1 moet zich aan de standaard bestelprocedures houden (dus alleen via secretariaat, DIA of CSC)

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Qari1 kan bestellingen rechtstreeks met het magazijn afhandelen

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Qarc2 moet weten waar precies een bestelling geplaatst moet worden (CSC of secretariaat/DIA)

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Qari2 kan met de eigen computer een bestelling indienen die vanzelf op de juiste plek komt

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Nu volgt een aantal uitspraken over de afhandeling van bestellingen in de mogelijk toekomstige situatie.
De medewerker moet…

Qari3 een statusoverzicht van de bestelling kunnen inzien op de eigen de computer

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Qarc4 zonder vooraankondiging de bestelling op de werkplek afgeleverd krijgen

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Qari4 een e-mail krijgen die de levering aankondigt

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Qarc3 voor het statusoverzicht van de bestelling bij secretariaat, Inkoop, DIA of CSC zijn

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

[17] The first page is missing because it contains strategic information about the organization.

De medewerker moet…

Qari5 na ontvangst van de bestelling per omgaande een e-mail versturen voor de ontvangstbevestiging

helemaal oneens	oneens	enigszins	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Qarc6 zelf navragen of er bij een bestelling iets mis is gegaan

helemaal oneens	oneens	enigszins	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Qarc5 ontvangstbevestiging geven door een handtekening op de afleverbon te zetten

helemaal oneens	oneens	enigszins	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Qari6 een waarschuwing per e-mail ontvangen als er bij een bestelling iets mis is gegaan

helemaal oneens	oneens	enigszins	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Nu volgt een aantal uitspraken over de 8m^2 gratis opslagruimte in het magazijn waarvan de diensten gebruik kunnen maken. Stel een dienst wil gebruik gaan maken van de 8m^2 gratis opslagruimte in het magazijn, bijvoorbeeld om drukwerk op te slaan.
In de mogelijk toekomstige situatie…

Qarc7 kan de medewerker de nog beschikbare ruimte slechts navragen bij het magazijnpersoneel

helemaal oneens	oneens	enigszins	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Qari7 wordt de nog beschikbare ruimte bijgehouden op ATRIUM

helemaal oneens	oneens	enigszins	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

SYSTEEMVEREISTEN GEKOPPELD AAN UW WERK

De volgende uitspraken betreffen een aantal eigenschappen van het nieuwe magazijnbeheersysteem en de mogelijke gevolgen voor de dagelijkse praktijk. Geef alstublieft aan in hoeverre u het met de uitspraken eens bent door een getal te omcirkelen 0 (helemaal oneens) - 5 (helemaal eens).

Qpvi-ppp1 Dat ik een statusoverzicht van de bestelling kan inzien op mijn eigen computer verbetert mijn controle op de bestellingen

helemaal oneens	oneens	enigszins	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Qpvi-npp2 Dat je je aan de standaard bestelprocedures moet houden (dus alleen via secretariaat, DIA of CSC) ondersteunt de flexibiliteit van de bestelprocedure

helemaal oneens	oneens	enigszins	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Qpvi-npp4 Dat je zelf moet navragen of er nog gratis magazijnruimte beschikbaar is, maakt de reservering van ruimte flexibel

helemaal oneens	oneens	enigszins	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Qpvc-nnn1 Dat je voor het statusoverzicht van de bestelling bij secretariaat, Inkoop, DIA of CSC moet zijn, is op een vervelende manier omslachtig

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Qpvi-ppn1 Dat je bestellingen rechtstreeks met het magazijn kan afhandelen, verkleint de kans dat een bestelling mislukt

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Qpvi-npp3 Dat je moet weten waar precies een bestelling geplaatst moet worden (CSC of secretariaat/DIA), bevordert de vlotte afhandeling van de bestelling

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Qpvc-nnp1 Dat je zelf moet navragen of er bij een bestelling iets mis is gegaan, hindert de accurate afhandeling van de bestelling

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Qpvi-npn3 Ontvangstbevestiging door een handtekening op de afleverbon te zetten, verkleint de kans op vergissingen

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Qpvc-nnn2 Dat een bestelling onaangekondigd op de werkplek wordt afgeleverd, maakt het werken chaotisch

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Qpvc-nnp4 Dat je zelf moet navragen of er nog gratis magazijnruimte beschikbaar is, vermindert de flexibiliteit van de reservering van ruimte

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Qpvc-nnp2 Dat je je aan de standaard bestelprocedures moet houden (dus alleen via secretariaat, DIA of CSC) vermindert de flexibiliteit van de bestelprocedure

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Qpvi-npp1 Dat je zelf moet navragen of er bij een bestelling iets mis is gegaan, bevordert een accurate afhandeling van de bestelling

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Qpvi-ppn3 Dat je zelf na ontvangst van de bestelling per omgaande e-mail de Ontvangstbevestiging geeft, vermindert de kans op fouten in de bestelprocedure

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Qpvi-ppn4 Het bijhouden van de nog beschikbare gratis magazijnruimte op ATRIUM voorkomt het nodeloos maken van kosten

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Qpvc-pnp2 Dat per e-mail de levering van een bestelling wordt aangekondigd, bemoeilijkt een goede planning

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Qpvc-pnn1 Dat je bestellingen rechtstreeks met het magazijn kan afhandelen,

vergroot de kans dat een bestelling mislukt

```
helemaal    oneens    enigszins    enigszins    eens      helemaal
oneens                oneens       eens                   eens
    0          1          2            3          4          5
```

Qpvi-npn2 Dat een bestelling onaangekondigd op de werkplek wordt afgeleverd, voorkomt dat het werken chaotisch wordt

```
helemaal    oneens    enigszins    enigszins    eens      helemaal
oneens                oneens       eens                   eens
    0          1          2            3          4          5
```

Qpvi-ppp3 Een waarschuwing per e-mail dat er iets mis is met de bestelling zorgt ervoor dat ik efficiënt kan werken

```
helemaal    oneens    enigszins    enigszins    eens      helemaal
oneens                oneens       eens                   eens
    0          1          2            3          4          5
```

Qpvc-pnp1 Dat ik een statusoverzicht van de bestelling kan inzien op mijn eigen computer verslechtert mijn controle op de bestellingen

```
helemaal    oneens    enigszins    enigszins    eens      helemaal
oneens                oneens       eens                   eens
    0          1          2            3          4          5
```

Qpvc-pnn2 Dat een bestelling via de eigen computer vanzelf op de juiste plek komt, maakt dat ik slordiger met de bestelprocedure omga dan voorheen

```
helemaal    oneens    enigszins    enigszins    eens      helemaal
oneens                oneens       eens                   eens
    0          1          2            3          4          5
```

Qpvi-npn1 Dat je voor het statusoverzicht van de bestelling bij secretariaat, Inkoop, DIA of CSC moet zijn, is op een prettige manier omslachtig

```
helemaal    oneens    enigszins    enigszins    eens      helemaal
oneens                oneens       eens                   eens
    0          1          2            3          4          5
```

Qpvc-pnp3 Een waarschuwing per e-mail dat er iets mis is met de bestelling maakt voor mij efficiënt werken moeilijk

```
helemaal    oneens    enigszins    enigszins    eens      helemaal
oneens                oneens       eens                   eens
    0          1          2            3          4          5
```

Qpvc-pnn3 Dat je zelf na ontvangst van de bestelling per omgaande e-mail de Ontvangstbevestiging geeft, verhoogt de kans op fouten in de bestelprocedure

```
helemaal    oneens    enigszins    enigszins    eens      helemaal
oneens                oneens       eens                   eens
    0          1          2            3          4          5
```

Qpvi-ppn2 Dat een bestelling via de eigen computer vanzelf op de juiste plek komt, maakt dat ik minder slordig met de bestelprocedure omga dan voorheen

```
helemaal    oneens    enigszins    enigszins    eens      helemaal
oneens                oneens       eens                   eens
    0          1          2            3          4          5
```

Qpvc-nnn3 Ontvangstbevestiging door een handtekening op de afleverbon te zetten vergroot de kans op vergissingen in de afhandeling van de bestelling

```
helemaal    oneens    enigszins    enigszins    eens      helemaal
oneens                oneens       eens                   eens
    0          1          2            3          4          5
```

Qpvi-ppp2 Dat per e-mail de levering van een bestelling wordt aangekondigd, vergemakkelijkt een goede planning

```
helemaal    oneens    enigszins    enigszins    eens      helemaal
oneens                oneens       eens                   eens
    0          1          2            3          4          5
```

Qpvc-nnp3 Dat je moet weten waar precies een bestelling geplaatst moet worden (CSC of secretariaat/DIA), stoort de vlotte afhandeling van de bestelling

```
helemaal    oneens    enigszins    enigszins    eens      helemaal
oneens                oneens       eens                   eens
    0          1          2            3          4          5
```

Qpvc-pnn4 Het bijhouden van de nog beschikbare gratis magazijnruimte op ATRIUM draagt bij aan het nodeloos maken van kosten

```
helemaal    oneens    enigszins    enigszins    eens      helemaal
oneens                oneens       eens                   eens
    0          1          2            3          4          5
```

GEBRUIKERSTEVREDENHEID HUIDIGE SYSTEEM

De volgende uitspraken betreffen de manier waarop op dit moment bestellingen worden afgehandeld en hoe tevreden u daarover bent. Geef alstublieft aan in hoeverre u het met de uitspraken eens bent door een getal te omcirkelen 0 (helemaal oneens) - 5 (helemaal eens).

Qcsi2 De flexibiliteit van de huidige manier van afhandelen van bestellingen door het Magazijn is goed

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Qcsc1 De manier waarop ik nu aan mijn bestellingen kom gaat voorbij aan mijn wensen

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Qcsc4 Automatische signalering dat de voorraad aanvulling behoeft, is voor mij nutteloos om tijd te besparen

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Qcsc3 De efficiëntie van de huidige manier van bestellingen doen via secretariaat, DIA of CSC is laag

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Qcsi1 De manier waarop ik nu aan mijn bestellingen kom voldoet aan mijn wensen

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Qcsc2 De flexibiliteit van de huidige manier van verwerken van bestellingen door het Magazijn is slecht

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Qcsi3 De efficiëntie van de huidige manier van bestellingen doen via secretariaat, DIA of CSC is hoog

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Qcsi4 Automatische signalering dat de voorraad aanvulling behoeft, levert mij een tijdsbesparing op

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

De volgende vragen hebben betrekking op uw bekendheid met de huidige mogelijkheden van opslag in het magazijn.

Enq1 Er is bij mijn dienst of sector iemand die over de $8m^2$ gratis opslagruimte in het magazijn gaat, waarbij ik voor vragen hierover terecht kan

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Enq2 Ik ben op de hoogte van deze mogelijkheid van gratis opslag in het magazijn

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Enq3 Ik ben op de hoogte van het feit dat je na deze 8m^2 gratis opslagruimte moet gaan betalen voor de gebruikte ruimte

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Enq4 Ik heb behoefte aan deze mogelijkheid van gratis opslag

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

Enq5 Ik ben op de hoogte van de procedure die ik moet doorlopen als ik gebruik wil maken van de gratis opslagruimte

helemaal oneens	oneens	enigszins oneens	enigszins eens	eens	helemaal eens
0	1	2	3	4	5

CR1 Uw op- en aanmerkingen:

DV1 Datum:
DV2 Organisatie: PU
DV3 Dienst:
DV4 Sector:
DV5 Functie:
DV5 Man/vrouw:
DV6 Leeftijd:
DV7 Aantal jaren in dienst:

Hartelijk dank voor het invullen van de vragenlijst. We hopen dat de resultaten zullen bijdragen aan een magazijn dat beter aangepast is aan uw behoeften.

Appendix 4.2

CHI RE*quest*

CHI RE*quest*

This questionnaire contains 24 statements concerning computer hardware, software, and functionality. While filling out this questionnaire, imagine you were to buy your own system and you could assemble it from the list of off-the-shelf products that this questionnaire offers you. The statements concern various considerations you might have while making your choice. You can indicate your level of agreement to the statements by checking a value on the respective rating scales (0= completely disagree, 5= completely agree). Your data will be processed anonymously and will be available to the researcher only.

---

pnn1 My system should have anti virus software with a monthly license fee

completely disagree	disagree	disagree a little	agree a little	agree	completely agree
0	1	2	3	4	5

pnp3 To get a 5 years pickup, repair, and return guarantee I am prepared to draw on my savings

completely disagree	disagree	disagree a little	agree a little	agree	completely agree
0	1	2	3	4	5

nnp2 My system should have a cathode ray tube monitor that emits high radiation into my eyes

completely disagree	disagree	disagree a little	agree a little	agree	completely agree
0	1	2	3	4	5

ppn3 I want a Linux operating system to prevent my system from becoming unstable

completely disagree	disagree	disagree a little	agree a little	agree	completely agree
0	1	2	3	4	5

---

npn3 I would like outdated browser software to relieve me from superfluous options

```
completely  disagree  disagree     agree       agree     completely
disagree              a little     a little              agree
    0          1          2          3          4          5
```

nnn3 My system should have a 1 GB hard disk to make me delete most of my files

```
completely  disagree  disagree     agree       agree     completely
disagree              a little     a little              agree
    0          1          2          3          4          5
```

ppp1 I want the latest AMD Athlon 64 processor so that I can work quickly

```
completely  disagree  disagree     agree       agree     completely
disagree              a little     a little              agree
    0          1          2          3          4          5
```

npp3 I'd like an old fashioned green-on-black display screen to improve the look and feel of my system

```
completely  disagree  disagree     agree       agree     completely
disagree              a little     a little              agree
    0          1          2          3          4          5
```

---

npp1 I opt for a second hand DOS machine to save me some cash

```
completely  disagree  disagree     agree       agree     completely
disagree              a little     a little              agree
    0          1          2          3          4          5
```

nnp1 I wish to have a stand alone computer to escape communicating with the international community

```
completely  disagree  disagree     agree       agree     completely
disagree              a little     a little              agree
    0          1          2          3          4          5
```

ppp3 I need a 350 GB hard disk to store most of my files

```
completely  disagree  disagree     agree       agree     completely
disagree              a little     a little              agree
    0          1          2          3          4          5
```

pnp2 I accept the risk of RSI in my hands to have a mouse device on my system

```
completely  disagree  disagree     agree       agree     completely
disagree              a little     a little              agree
    0          1          2          3          4          5
```

---

nnn2 To get a 5¼" floppy drive I want to increase my costs

```
completely  disagree  disagree     agree       agree     completely
disagree              a little     a little              agree
    0          1          2          3          4          5
```

npp2 I want to work only from the prompt to enhance my working pace

```
completely  disagree  disagree     agree       agree     completely
disagree              a little     a little              agree
```

```
    0          1          2          3          4          5
```

ppn1 I need a connection to the Internet to avoid working in isolation

```
completely  disagree  disagree     agree       agree     completely
disagree              a little     a little              agree
    0          1          2          3          4          5
```

npn2 The installation wizard should skip all help and tutorials to prevent me from becoming lazy

```
completely  disagree  disagree     agree       agree     completely
disagree              a little     a little              agree
    0          1          2          3          4          5
```

---

ppp2 I want a well-designed interface to increase the usability of my system

```
completely  disagree  disagree     agree       agree     completely
disagree              a little     a little              agree
    0          1          2          3          4          5
```

pnp1 To get a 63” Wide Screen Plasma Monitor I am willing to drain my budget

```
completely  disagree  disagree     agree       agree     completely
disagree              a little     a little              agree
    0          1          2          3          4          5
```

nnp3 To get a Windows ’95 operating system I am ready to risk the stability of my system

```
completely  disagree  disagree     agree       agree     completely
disagree              a little     a little              agree
    0          1          2          3          4          5
```

nnn1 My system should have an old 486 DX processor to slow me down

```
completely  disagree  disagree     agree       agree     completely
disagree              a little     a little              agree
    0          1          2          3          4          5
```

---

npn1 I want big manuals that I should read first to keep me from making errors

```
completely  disagree  disagree     agree       agree     completely
disagree              a little     a little              agree
    0          1          2          3          4          5
```

pnn3 I am ready to pay extra for a Laser All-in-One printer, copier, fax machine, and scanner

```
completely  disagree  disagree     agree       agree     completely
disagree              a little     a little              agree
    0          1          2          3          4          5
```

pnn2 My system should have such an outstanding firewall that it challenges hackers to attack

```
completely  disagree  disagree     agree       agree     completely
disagree              a little     a little              agree
    0          1          2          3          4          5
```

ppn2 I want a TFT monitor to reduce possible damage to my eyes

```
completely  disagree  disagree     agree       agree     completely
```

```
disagree              a little     a little              agree
    0          1          2          3          4          5
```

DV1 Age:

DV2 Male/female:

DV3 Function:

DV4 Organization where you work or study:

DV5 Department:

DV6 Section:

DV7 Additional remarks:

Appendix 4.3

VT*Player* RE*quest*[18]

De Braillemuis is belangrijk voor mijn

leren
concentratie
begrip
fantasie
zelfstandigheid
leerplezier
inzicht
werklust
cijfers
stemming

PPP-items
PPP1 Rechte lijnen zijn makkelijk te volgen
PPP3 Ik begrijp het plaatje beter als ik alleen de belangrijkste onderdelen krijg
PPP4 Geluidseffecten verhelderen de betekenis van onderdelen
PPP6 Aan de lijndikte kan je zien wat belangrijk is
PPP7 Een gesproken inleiding bij het plaatje helpt bij het leren
PPP8 Ik krijg meer inzicht als de stem iets uitlegt
PPP9 Als ik op onderdelen moet klikken, is het duidelijk bij welk onderdeel de informatie hoort
PPP12 Het plaatje is helder zonder details binnenin de vlakken

PNP-items
PNP1 Rechte lijnen zijn moeilijk te volgen
PNP5 De verschillende arceringen verhinderen dat ik vlakken uit elkaar kan houden
PNP8 Ik krijg minder inzicht als de stem iets uitlegt
PNP12 Het plaatje is ingewikkeld zonder details binnenin de vlakken

PPN-items
PPN4 Geluidseffecten verhinderen dat er teveel informatie op me af komt
PPN6 De verschillende lijndiktes verhinderen dat ik me in details verlies
PPN7 Een gesproken inleiding bij het plaatje beperkt mijn verwarring
PPN8 De stem voorkomt dat ik mijn concentratie verlies
PPN10 Bij een grote afbeelding is de kans klein dat ik me vergis in mijn positie
PPN11 Bij stapsgewijze begeleiding is hulp overbodig

PNN-items
PNN1 Door rechte lijnen raak ik de weg kwijt
PNN2 Als vormen simpel zijn gaan ze op elkaar lijken
PNN3 Als ik alleen de belangrijkste onderdelen krijg raak ik het overzicht kwijt
PNN5 Door de verschillende arceringen haal ik vlakken door elkaar
PNN9 Als ik op onderdelen moet klikken is informatie moeilijk vindbaar
PNN10 Ik vergis me in mijn positie als de afbeelding groot is
PNN12 Het verwart me als details binnen de vlakken ontbreken

NPP-items
NPP1 Kromme lijnen zijn makkelijk te volgen
NPP2 Ingewikkelde vormen zijn goed te begrijpen
NPP3 Ik begrijp het plaatje beter als ik ook de bijzaken krijg
NPP4 Verkenning zonder geluidseffecten maakt het plaatje duidelijk
NPP5 Dat dezelfde arcering een paar keer wordt gebruikt helpt me de vlakken uit elkaar te houden
NPP6 Als alle lijnen even dik zijn kan ik goed zien wat belangrijk is
NPP7 Zonder informatie vooraf kan ik beter leren
NPP8 Ik krijg meer inzicht via de braille leesregel
NPP10 Bij een klein plaatje houd ik de vlakken makkelijk uit elkaar
NPP12 Lijnen binnenin de vlakken maken het plaatje helder

[18] Transcription of the audio files.

NNP-items
NNP1 Kromme lijnen zijn moeilijk te volgen
NNP2 Ingewikkelde vormen zijn slecht te begrijpen
NNP9 Als informatie automatisch wordt afgespeeld twijfel ik over welk onderdeel er verteld wordt

NPN-items
NPN1 Kromme lijnen voorkomen dat ik de weg kwijt raak
NPN2 Ingewikkelde vormen verhinderen dat onderdelen op elkaar gaan lijken
NPN3 Dat ik ook de bijzaken krijg voorkomt dat ik het overzicht kwijtraak
NPN7 Beginnen zonder informatie vooraf vermindert mijn twijfel over het onderwerp
NPN8 Lezen via de braille leesregel beperkt mijn concentratieverlies
NPN11 Als ik het plaatje zelfstandig moet verkennen is hulp overbodig

NNN-items
NNN1 Kromme lijnen verwarren me
NNN2 Als vormen ingewikkeld zijn gaan ze op elkaar lijken
NNN5 Ik raak in de war als ik steeds dezelfde arceringen tegenkom
NNN8 Ik verlies mijn concentratie als ik informatie lees via de braille leesregel

5 Usability and Efficiency

Abstract

Although it seems that the field has moved beyond usability taken as mere performance, concepts such as effectiveness, efficiency, and effort have not been worked out properly.[1] In industry as well as academia, the ISO 9241-11 norm provides the dominant view on usability, stating that usability is a function of effectiveness, efficiency, and satisfaction. Although intuitively, usability requirements should be part of a software's design in an early stage, conceptually and empirically, it seems more likely that performance requirements (i.e. the absence of errors) should be the center of concern. This chapter offers an elaborated view on usability, satisfaction, and performance. The theoretical conceptions are tested with data gathered from users of banking and hospital systems by means of a 4-years single-item survey and a structured questionnaire, respectively. Results suggested that performance factors (i.e. efficiency) are more important than usability in understanding why stakeholders are satisfied with a system or not. Moreover, it neither is dissatisfaction with a system nor that a system is less usable that predicate requirements change. Instead, avoiding machine inaccuracy best predicted the variability in agreement to must requirements, while achieving human accuracy predicted the variability in agreement to the won't requirements. The chapter closes with the argument that good programming beats proper interface design.

Keywords: Stakeholder goals, Effectiveness, Efficiency, Effort, Performance, Usability, Satisfaction.

5.1 Introduction

"I want to file a complaint. To reload your e-commerce pages, I need to leave the dialog box and look up the link in the bookmarks list. It takes me more time than necessary and I often make mistakes. I find this very inconvenient." The problem of reloading Web pages through the bookmarks list is not a matter of sloppy programming but rather of poor usability. It is a matter of overlooking operability in the early stages of system design. And although the site may be free of computational defects, it may still arouse the dissatisfaction of

[1] This chapter is partly based on Hoorn (2005a).

its customers when usability issues are ignored during requirements gathering and analysis.

To keep the customers satisfied, so the argument goes, the definition of a system's usability should be one of the central activities during requirements development (Jokela et al., 2003). According to ISO 9126 (1991), usability is one of the quality requirements of a system just like reliability and maintainability are. However, usability as a concept is confused (Van Welie et al., 1999), inconsistently used (Jokela et al., 2003), and may vary for different stakeholder groups, working in different contexts of use (Jokela et al., 2003). Nevertheless, it is important to get a grip on what aspects contribute to usability so that we better understand on what grounds stakeholders are willing to learn, operate, and accept the interactive systems we develop.

Several attempts have been made to analyze the concept of usability (e.g., ISO 9241-11 (1998); Van Welie et al., 1999; Jokela et al., 2003). This chapter is just another try to push our understanding a little further. Almost every author admits that usability is a multi-faceted concept (e.g., Jokela et al., 2003). This chapter will not attempt, therefore, to treat every aspect that may relate to usability. The focus is on usability as a prerequisite of user or stakeholder satisfaction, while usability itself is determined by, among other factors, effectiveness and efficiency. At times, major changes are needed in systems that are already up and running, which is a time-consuming and costly endeavor. The purpose of the conceptual analysis presented in this chapter and its empirical verification with users of banking and hospital systems is to bring more focus to requirements activities by discussing and testing those usability aspects that may deeply contribute to stakeholder satisfaction. Another aim of this study, then, is to make usability and related concepts more measurable (Nielsen & Levy, 1994). Measurable usability requirements are important because "what is measured gets to be done" (Jokela et al., 2003).

5.2 Usability as a Concept

Usability as a concept typically is a concern of interaction designers and CHI-experts. This section provides a quick overview of what the community thinks usability is about (for an overview of standards, see Bevan, 2001). It then turns to a selection of usability dimensions to create a focal point of empirical verification.

Central to almost all usability discussions are the ISO 9241-11 norms, featuring the citation that usability is "the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction." This 'standard' definition of usability (Jokela et al., 2003) is adopted also by the Common Industry Format for usability testing ANSI/INCITS 354 (2001). Many authors mention the triplet of effectiveness, efficiency, and satisfaction in alliance with some extras such as learnability, customizability, and helpfulness. For example, Shackel (1991, p. 25) defines usability not only in terms of effectiveness but also as learnability, flexibility,

and attitude. Nielsen mentions, apart from efficiency and satisfaction, the terms learnability, memorability, and errors/safety (Nielsen, 1993). Newman and Lamming mention satisfaction, learning, memorability ("retention of learned skills"), and errors, but they add performance speed, customization, and reorganization activities (Newman & Lamming, 1995, pp. 30-31). In Shneiderman's view, the five factors satisfaction, errors, (time to) learn, memorability (retention over time), and performance speed are most germane (Shneiderman, 1998). Another categorization comes from Dix et al. (1998), who consider flexibility, robustness, and again, learnability to be of central importance. Jordan (1998) defines usability not only as efficiency but also as guessability, learnability, and re-usability. Wickens et al. (p. 2) consider performance, safety, and satisfaction the most important factors (Wickens et al., 2004). Lauesen (p. 9) mentions as usability factors satisfaction, efficiency, learning, memorability, and understandability (Lauesen, 2005). Apart from error management and adaptability (customization), Scapin and Bastien (1997) emphasize requirements such as guidance, control, consistency, significance of codes, and compatibility. Most importantly, however, Scapin and Bastien explicitly mention workload as a usability factor, which for the other authors remains implicit in items such as 'ease of learning' or 'ease of remembering' (e.g., Lauesen, 2005, p. 9).

In a most enlightening review, Van Welie et al. analyze, compare, and integrate the concepts of ISO 9241-11, Shneiderman, Nielsen, Scapin and Bastien, and Dix et al. to arrive at a model of usability (Van Welie et al., 1999) instead of a set of bulleted lists. For example, they argue that Dix's robustness can be seen as an instantiation of effectiveness. Moreover, Van Welie et al. clarify that the various authors approach the concept of usability at different but related levels of abstraction. At the top level is the standard ISO definition, stating that usability depends on a system's effectiveness, efficiency, and the user's satisfaction. The second layer consists of so-called usage indicators, which are variables observable at the work floor, such as performance speed, errors, learnability, and memorability.[2] For example, performance speed at the work floor is an indicator of the level of overall efficiency. Underlying the usage indicators are the means (the model's third layer) to achieve performance speed, learnability, etc. The model claims, for example, that consistency and adaptability together feed performance speed. The fourth layer consists of knowledge domains (e.g., design knowledge) that can be used to improve the means but these variables need not concern us here. For now, the distinction between usability factors and usage indicators is most useful to distinguish the explanation (i.e. usage indicators) from what should be explained (e.g., effectiveness and efficiency). Yet, the means are important in their own right. They more-or-less represent the requirements (e.g., consistency, adaptability) and

[2] It remains to be seen whether learnability and memorability are readily observable in practice.

system specifications (e.g., feedback, warnings, shortcuts, undo) that help establish the stakeholder's impression that the system is error free, safe, or learnable (Van Welie et al., 1999). However, for the conceptual analysis the means/requirements are less useful. One needs to understand the concept before designing the handle.

Much in line with the previously mentioned authors, the Preece's text book (pp. 18-19) also distinguishes effectiveness, efficiency, learning, memorability ("easy to remember how to use"), safety, and "good utility" as usability goals (Preece et al., 2002). However, these authors clearly separate the usability goals from user experiences such as 'aesthetically pleasing' and 'fun' but most importantly, from satisfaction. Preece and her colleagues, then, implicitly state that the relation of effectiveness and efficiency with usability differs from that of satisfaction with usability whereas the standard ISO definition and her supporters do not make such a distinction. The next section elaborates on this matter since the separation of effectiveness and efficiency from satisfaction will turn out to be crucial in understanding the status of usability.

5.3 Usability and its Relation to Satisfaction

In the usability description of ISO 9241-11, the role of satisfaction is somewhat obscure. It is not too clear whether satisfaction feeds the level of usability or usability feeds the level of satisfaction. In the first case, satisfaction forms the input for usability just like effectiveness and efficiency do. In the second case, satisfaction is the net-result of an effective, efficient, and usable system. Authors such as Brooke et al. (1990) and Frøkjær et al. (2000) conjecture that satisfaction explains usability. Simply put, these authors claim that what makes a person happy is something usable. Yet, all things that make a person happy are not *necessarily* useful things (e.g., an aesthetically pleasing color) but all useful things (e.g., a corkscrew) *possibly* make people happy. This means that stakeholder satisfaction is not only established through the usability of a device but also through, for example, aesthetics, fun, and triggering creativity (Preece et al., 2002, p. 19).

In addition, satisfaction is not only used to estimate usability. It is also used to indicate effectiveness, for instance, of information systems (Pather et al., 2003). Contrariwise, Scott (1995) reviews literature claiming the opposite that effectiveness of an information system is responsible for user satisfaction. With Preece et al. (2002), I believe that satisfaction does not have the same status as effectiveness and efficiency. In this chapter, I assert that satisfaction is elicited *after* goal achievement, whereas usability and use relate to task execution while achieving such goals. Satisfaction, then, is the end product of interacting with a usable system and not a stage in between. Later in the chapter I will present empirical evidence to corroborate this assumption.

5.4 Usability and its Relation to Performance

If satisfaction is the end product and usability its main nutrient, then the two performance factors of effectiveness and efficiency in their turn foster usability. That makes usability something in between performance and satisfaction. Love (1991) states that effectiveness and efficiency are two crucial aspects of performance (also Nielsen & Levy, 1994; Jordan, 1998). Wickens et al. (2004, p. 2) consider performance an "all-encompassing term" that involves productivity (an aspect of effectiveness, see Barnard, 1938) as well as error reduction and production speed (in my view, these two together make up efficiency). With regard to effectiveness, most authors follow Barnard (1938), who states that effectiveness indicates in how far the actual outputs of a system correspond to the desired outputs. That is, in how far a stakeholder achieves a desired goal with the system (ISO 9241-11, 1998).

Yet, variables preeminently mentioned as performance factors are "time to complete a task (efficiency)" and 'the number of errors during task execution' (Preece et al., 2002, p. 18; Newman & Lamming, 1995, pp. 30-31). Learnability and memorability are also mentioned in this respect, in particular in relation to the error rate of task execution, but these are instantiations of the more general aspect of response times in combination with error rate as indicators of efficiency. Efficiency, for quite a few authors, refers mainly to time aspects of process execution or task completion (e.g., Frøkjær et al., 2000; ISO 9241-11). However, others claim that speed per se is meaningless unless it is combined with levels of accuracy (Käki, 2004; Ulrich & Hebert, 1982).

It seems, then, that the main performance factor is speed or time-efficiency, followed by the reduction of errors (sometimes seen as an aspect of efficiency, sometimes as effectiveness). Effectiveness seen as the achievement of stakeholder goals (e.g., increased productivity) is mentioned every now and then but what really is missing out as a performance factor is the effort a user invests in handling an interactive system. This is strange, because cognitive ergonomics (e.g., Scapin & Bastien, 1997) and Human Factors (e.g., Wickens et al., 2004, p. 149) put a lot of emphasis on workload reduction, systems being easily operable, easy to learn, easy to remember, easy to maintain, while avoiding information overload for the user. Effort relates to the amount of labor stakeholders put into handling a system (Cooper, 1968; Silverstein et al., 1998) but can also refer to cognitive load (Oviatt et al., 2004). It seems likely that effort is the third performance factor because in particular efficiency is negatively affected by an increase in task-difficulty (Oviatt et al., 2004). Yet, although it is widely acknowledged that performance as reflected by effectiveness and efficiency contributes considerably to usability (e.g., ISO 9241-11; Nielsen & Levy, 1994), effort is mentioned only indirectly (e.g., the 3-click rule) as a predictor of usability (Cooper, 1968). Wickens et al. warn designers that a powerful interface feature "will go unused" if the effort costs are high despite the anticipated gains in effectiveness (i.e. productivity) (Wickens et al.,

2004, p. 149). In this study, therefore, I treat effort together with effectiveness and efficiency as one of the main performance factors that affect usability, and via usability, user satisfaction.

5.5 Stakeholder Logistics

Figure 1 offers a summary of the theoretical assumptions made so far and an elaboration of the dependency among the performance factors. In Figure 1, Usability is modeled as a threefold function of Effectiveness, Efficiency, and Effort. Love (1991) states that effectiveness and efficiency are *independent* variables. In following ISO 9241-11, Frøkjær et al. (2000) correlated effectiveness, efficiency, and satisfaction with usability but did not establish significant results. These authors yet posit that effectiveness, efficiency, and satisfaction are directly responsible for the level of usability of a system. In Figure 1, then, Effectiveness and Efficiency are modeled as independent predictors of Usability and for the sake of simplicity; Effort is modeled in the same way. For reasons explained in Section 5.3, Satisfaction is not modeled as a predictor or mediator but as a dependent measure, being a judgment that covers all kinds of aspects of a system – not usability alone. The model recognizes that other factors (dashed) can contribute to Usability (e.g., learnability) and to Satisfaction (e.g., aesthetics) but these fall outside the scope of the present chapter.

In opposition with Love (1991) and Frøkjær et al. (2000), Oviatt et al. (2004) argue that the three performance factors are *dependent*. Oviatt et al. (2004) manipulated the levels of task difficulty and found that the number of task-critical errors and response latencies increased significantly when the task became more difficult. Thus, it seems that efficiency as a function of speed and accuracy is related to the degree of cognitive processing load or effort (also Wickens et al., 2004, p. 149). In Figure 1, then, assuming partial independence is the middle way between Love and Oviatt et al. The three Performance factors are supposed to correlate (double-headed arrows) so that, for example, an increase in Effort can be negatively related to Efficiency or an increase in Efficiency is positively related to high Effectiveness.

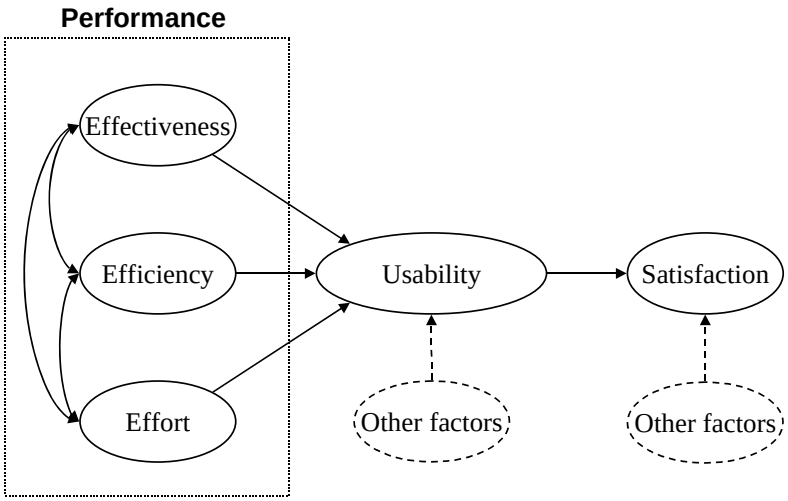


Figure 1: Model of Stakeholder Logistics. Usability of an interactive system mediates between Satisfaction on the one hand and Performance (Effectiveness, Efficiency, and Effort) on the other.

5.6 The Human and the Machine Side of Performance

In this section, I will elaborate the three performance factors for both their machine and their human side. After all, both humans and machines are making up the 'interactive system' and both parties can contribute to making errors, being quick, and achieving goals. Judgments on humans and machines rely on perceptions of the physical and directly observable variables of a system (cf. usage indicators Van Welie et al., 1999), such as the number of target hits or clock time of process execution. This means that if a system is considered 'fast,' this judgment is based on the perception of the physical clock time to execute a task to which the stakeholder, moreover, attributes a subjective weight. To satisfy the stakeholders of an interactive system, the high-level requirement of usability is fostered by three lower-level performance requirements. Together, these performance requirements constitute the best possible stakeholder logistics a system could offer. These requirements relate to the desired state that the three performance factors should be in. That is, *an interactive system should enable the largest possible degree of goal accomplishment (effectiveness), as fast as possible and against the smallest possible number of errors (efficiency), at an optimal (which need not be minimal) level of effort.* These performance requirements also govern matters such as learnability and memorability because stakeholders probably want to learn and remember as much as possible, as fast as possible while making the smallest possible number of mistakes against some optimal (which could be minimal) level of effort.

5.6.1 *Effectiveness (output)*

In following Drucker (1954; 1974), Love (1991) asserts that effectiveness is related to goal end states ("doing the right things"), whereas efficiency is process oriented ("doing things right"). Frøkjær et al. (2000) envision effectiveness as "the outcome of the user's interaction with the system." Interpreted within Fitts' (1954) paradigm, effectiveness would be a reflection of the number of times the user hits a target in contrast to the number of misses. However, judgments on effectiveness are not bound to low-level interaction issues but can also concern achieving business goals and other higher objectives with a system (e.g., Hamilton & Cervany, 1981). In that case, effectiveness can also be a function of hitting and missing a number of business targets.

Effectiveness of an interactive system, then, is related to the degree that a stakeholder achieves a goal with the system. This goal may be related to the computer task at hand (e.g., to find a word in a document) or something outside that particular computer task (e.g., being entertained by the animated agent of a help function). Stakeholders assess effectiveness in terms of result (Love, 1991) and success (Seddon et al., 1998). Seddon et al. (1998) and Novick (1997) make an important distinction between the effectiveness of humans and the effectiveness of machines in goal accomplishment. Thus, effectiveness can be measured by the degree that a stakeholder perceives or experiences that a goal is achieved on the one hand by the machine and on the other by humans. Envisioning the stakeholder logistics as an assembly line, effectiveness is the result of successfully putting together the end product (cf. Barnard, 1938).

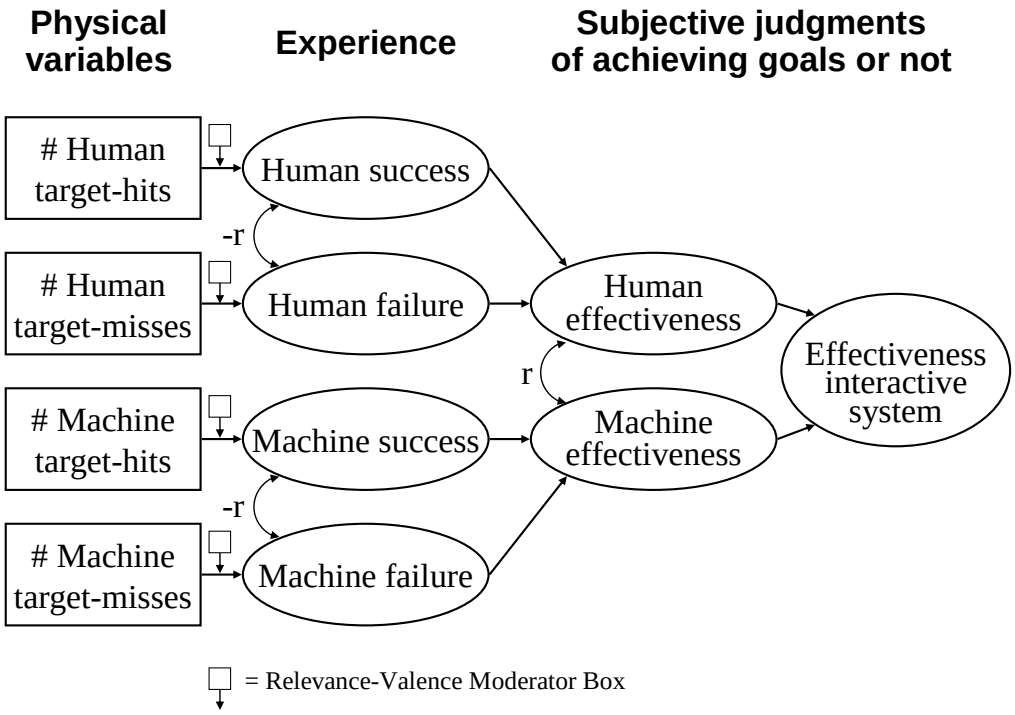


Figure 2: Explanation of judgments on effectiveness of an interactive system.

In Figure 2, Effectiveness of an interactive system is modeled not as an aspect inherent in a system but as a judgment (whether by experts or not) about that system. These judgments are fed by physically countable variables, such as the number of target hits and misses by both human and machine. Stakeholders experience the relative number of hits and misses in terms of success and failure, respectively. Common sense would have it that a ratio or percent expresses the relation between the numbers of hits versus misses. However, from a stakeholder's point of view, this is not necessarily the case. It is not hard to imagine that certain stakeholders are positively biased to their own success and underestimate their own degree of failure. In addition, power users may attribute machine success to their own doings and their own human failure to the machine. Novices or low-achievers may do the opposite, thinking they are too clumsy to handle a complex machine. In other words, stakeholders put different weights or relevance to the number of hits and misses so that the relation between hits and misses can differ from conventional ratio. Therefore, the experience of success and failure are modeled in Figure 2 as two relatively dependent variables, their interrelation being expressed by a negative correlation (-r) (Figure 2, double-headed arrows).

Figure 2 further indicates that the experience of success and failure precedes the more formal and reflective judgments of how effective (effective vs. ineffective) the human aspect of interaction was in hindsight. The same occurs

for the machine aspect (effective vs. ineffective). Between human effectiveness and that of the machine, a positive correlation is hypothesized, pointing out that an increase in effectiveness of the machine probably co-occurs with higher effectiveness of the human and v.v.

Finally, the judgment of human effectiveness and the judgment of machine effectiveness are combined into one overall judgment of the Effectiveness (effective vs. ineffective) of the interactive system. In its turn, the overall Effectiveness of the interactive system (humans in cooperation with machines) explains part of the variance of Usability, which explains part of the variance of Satisfaction (Figure 1).

Taking a closer look at Figure 2 reveals that between the physically countable variables and experience, the effects of two moderators occur (♆): Relevance (Chapter 3) and Valence (Chapter 4). That is, stakeholders estimate the relative importance (Relevance) of the physical variable to their personal or business goals and concerns and estimate in how far this damages or supports those goals and concerns (Valence). Only then can experience take place and do emotions occur (Frijda, 1986). A more detailed account of the Relevance-Valence Moderator Box is provided in Section 5.7. In the next section, I will argue that Effectiveness is relatively dependent on but not similar to Efficiency.

5.6.2 *Efficiency (throughput)*

Although in the literature, the definition of effectiveness is somewhat confused, compared to efficiency the discussion seems to be transparent. In quite a few cases, aspects of effectiveness are attributed to efficiency and the other way round or efficiency is treated as a mere aspect of time.

To start with the ISO 9241-11 (1998) norm, effectiveness is defined as the accuracy and completeness of goal achievement. ISO 9241-11 regards error rates as one of the indicators of effectiveness (also Nielsen & Levy, 1994) and I suppose as related to the accuracy aspect. Efficiency, so the ISO norm runs, pertains to the *relation* between accuracy and completeness of goal achievement. Moreover, time to execute a task is regarded a stable indicator of efficiency (also Nielsen & Levy, 1994). Frøkjær et al. (2000) follow the ISO definition of efficiency in focusing on the time aspect alone.

In view of the previous section, there will be not much discussion that completeness of end-goal achievement is the core of effectiveness. However, accuracy of goal achievement seems to be mistaken. An end goal can be completely achieved without being achieved in an accurate way (just being lucky). A shot can completely miss the target although the accuracy of aiming the gun was perfect (bad luck, target moved away). In other words, accuracy seems to be an aspect of the process towards end-goal achievement rather than an aspect of end-goal achievement in itself. Put differently, “it is quite possible for a manager to work efficiently and still remain ineffective” (Wambugu, 1982). In support of that, Love (1991) states that “even the most efficient organiza-

tion cannot survive if it is efficient at doing the wrong things. Likewise, the organization with the greatest effectiveness can disintegrate from poor efficiency."

Thus, if accuracy does not belong to end-goal achievement but to process execution then it is not an aspect of effectiveness but rather of efficiency and so are the related error rates. Making errors is missing the sub goals while striving for the end goals, that is, being effective. What remains then from the ISO definition of efficiency is the time aspect of process execution. Yet, as ISO 9241-11 (1998) states, there can be a *relation* between accuracy (of the process) and end-goal achievement and I will return to this matter later.

From the previous paragraphs it can be learned that efficiency relates to task or process execution and that it has an accuracy and a time aspect. Efficiency is subordinate to effectiveness because it is just a means to an end. The experiential interpretation of the time to execute a process or to complete a task is speed ("My system is fast, your system is slow"), which is relative to previous experiences. The Usability Glossary (2005) states:

> When asked to perform a task as well as possible, people will apply various strategies that may optimize speed, optimize accuracy, or combine the two. For this reason, comparing the performance of 2 users cannot be done on the basis of speed or accuracy alone, but both values need to be known. (Usability Glossary, 2005)

To be meaningful, hence, speed should be taken together with accuracy (i.e. error rates) for humans (Käki, 2004) as well as for machines (Ulrich & Hebert, 1982). In other words, process execution is governed by the well-known phenomenon that increases in speed are traded for decreases in accuracy and v.v (Usability Glossary, 2005; Oviatt et al., 2004). Efficiency then is one of four possible states of a speed-accuracy trade-off that a process can be in. These states are fast-accurate, fast-inaccurate, slow-accurate, slow-inaccurate. The first state is commonly referred to as 'efficient,' whereas the latter is deemed 'inefficient.' The two remaining combinations are sub-optimal states of efficiency that a process can be in. This constellation can apply to perceptual-motor and other human tasks as well as to running a process on a machine.

Thus, efficiency of an interactive system has little to do with the goal in itself as it is related to the means or processes to achieve those goals. The relation between accuracy and end-goal achievement that ISO 9241-11 hinges on should be seen as follows. Achieving a desired goal or not may depend on the accuracy of the associated process. Accuracy has to do with achieving certain sub goals that lead to achieving the main goal. However, not every sub goal has to be reached in order to get to the main goal. In other words, there is a relation between accuracy and the completeness of goal achievement but the first is not necessary to achieve the other (e.g., Wambugu, 1982). The only

goal that should be accurately hit for the system to be called effective is the end goal.[3]

Efficiency should be decomposed into a time (i.e. speed) aspect and an accuracy aspect. The real-time a process takes to achieve a goal with the system (e.g., connecting to the Internet) can be experienced as fast or slow. Time-experience, then, is one of two components of efficiency. Accuracy, whether the real number of errors humans and machines make is experienced as error prone or not, is the other. If connecting to the Internet is fast but at the cost of making many errors, the process is deemed – at least partially – inefficient. If the user accurately follows a wizard but at the cost of being extremely slow, the wizard also seems (partially) inefficient.

My position is that the degree of estimated efficiency is a function of the experience of two speed-accuracy trade-offs that each can be in one of four states (Figure 3). Process execution by humans (Käki, 2004) or machines (Ulrich & Hebert, 1982) can be experienced as slow and inaccurate, slow but accurate, fast but inaccurate, or fast and accurate. The latter combination is usually considered the optimal state of a system's processes but does not need to be in all cases. For example, in their meta-analysis, Nielsen & Levy (1994) found that the highest speeds in process execution did not necessarily lead to the highest levels of satisfaction. Using the analogy of the assembly line, efficiency is a function of the speed of the assembly belt and the precision with which, for instance, all the different parts of a computer chip are mounted on a surface.

Efficiency of an interactive system relates to the supply, distribution, and exchange of information within and between humans and machines. Figure 3 displays the constellation of speed-accuracy trade-offs in process execution of both humans and machines as based on clock time and precision. Time is a continuous variable and therefore one would expect that the experience of process time is either "fast" or "slow." Figure 3, however, shows that the same time epoch of a (stage in the) process can be experienced as fast as well as slow. This has to do with the interference of Relevance and Valence (Section 5.7). Suppose connecting to the Internet from a remote location is urgent (Relevance is high) to check an important message in the e-mail box. Suppose this message tells the reader whether s/he is accepted for a tenure track position or not. Then the Valence towards the message can be positive (expecting acceptance) or negative (expecting rejection). If the applicant anticipates acceptance, connecting to the Internet will be appraised as slow rather than fast ("I can't wait"). Conversely, if the applicant fears to be rejected, the same time to establish the connection is experienced as (too) fast. If the applicant is in doubt about the outcome of the message, parallel experiences of fast as well as slow will occur. This evaluative inconsistency becomes even clearer when the

[3] This explains why Dutch football is highly efficient yet ineffective. The players play brilliantly but often miss the goal. Nice to look at but not if you are in it for the championship.

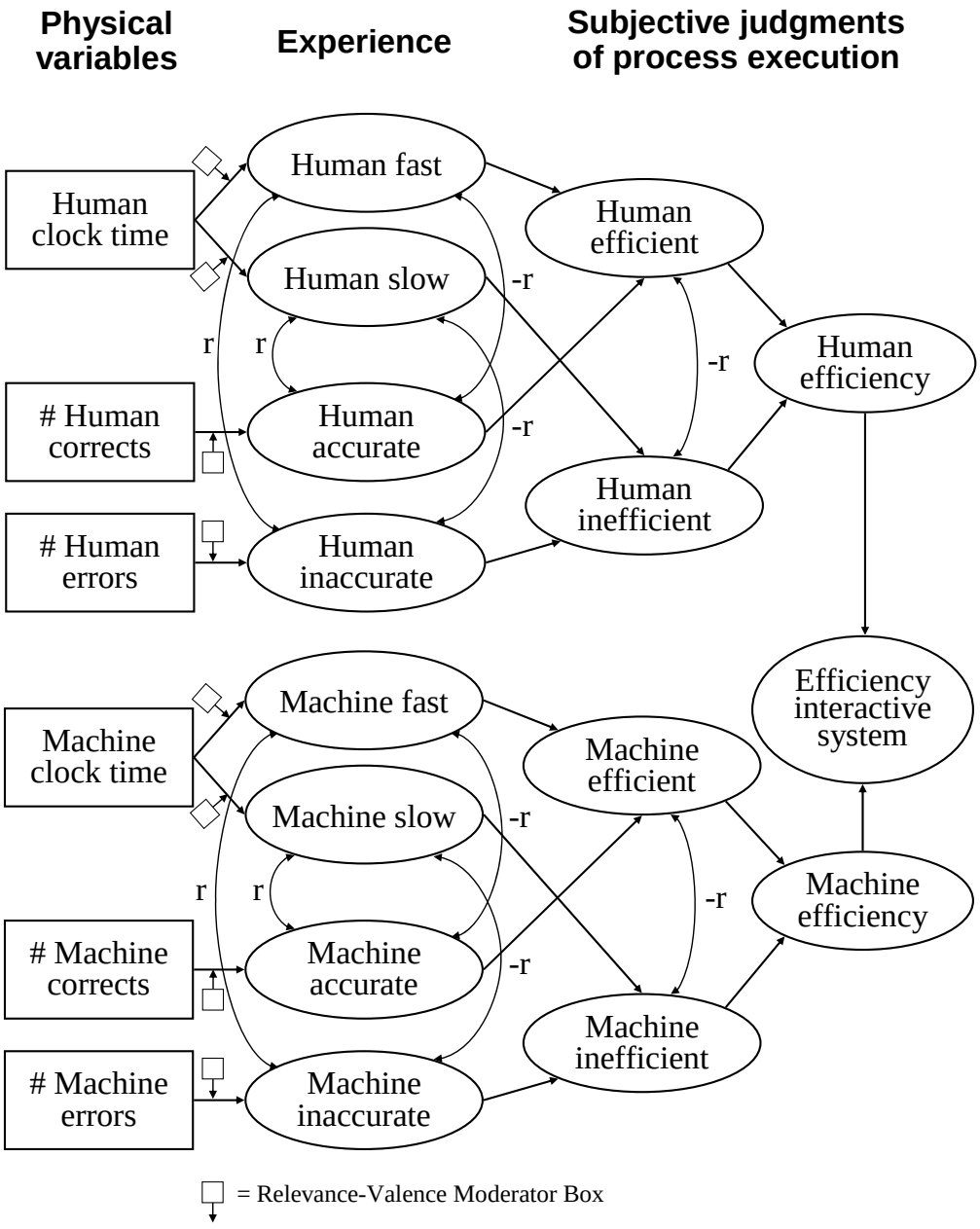


Figure 3: Explanation of judgments on efficiency of an interactive system.

time to reboot after a failure (machine experienced as slow) equals the time a virus needs to erase a disk (too soon).

The other two physical variables that foster efficiency judgments are the number of correctly executed stages and procedures in a process in some sort of negative relation (-r) to the number of errors. This situation resembles the one described for the numbers of target hits and misses (Section 5.6.1). The relation between correct process-stage execution and errors logically would be a percentage but psychologically, estimated accuracy may not be fully related to estimated inaccuracy.

The speed-accuracy trade-offs that occur at the level of physical variables are reflected in the stakeholder's experience as a set of four correlations for humans and four for machines (Figure 3). That a process is appraised as fast is positively related to the experience of inaccuracy (faster and with more imprecision) and negatively to accuracy (faster and with less precision). That a process is experienced as slow is positively related to accuracy (slower and with more precision) and negatively to inaccuracy (slower and with less imprecision). Experiences of "fast" in combination with "accurate" lead to the judgment of efficiency. Experiences of "slow" in combination with "inaccurate" lead to the judgment of inefficiency. Mixed combinations are also possible. Processes that are regarded fast as well as inaccurate (e.g., quick and dirty methods) will be partially considered efficient as well as partially inefficient. Processes that are regarded slow as well as accurate (e.g., check-and-double-check security policies) will show a similar ambivalence in the efficiency judgments. Subsequently, the combined judgments on human and machine efficiency and inefficiency predicate the level of overall Efficiency (efficient vs. inefficient) that is attributed to the interactive system as a whole. This judgment should explain a large degree of the system's Usability, which is responsible for a significant part of the variance in Satisfaction (Figure 1).

5.6.3 *Effort (input)*

In discussions on effectiveness, efficiency, and usability of interactive systems, effort seems to be out of the picture or is treated separately under the heading of ergonomics (e.g., Scapin & Bastien, 1997; Wickens et al., 2004, p. 149). In Barnard's (1938) vision on efficiency, certain aspects of effort are mixed in with aspects of effectiveness. Barnard characterizes efficiency as the ratio of completed outputs to actual inputs, such as money, skills, or workforce. ISO 9241-11 (1998) follows the same line of thought in that efficiency is not only the "accuracy and completeness with which users achieve certain goals" but also the "resources expended in achieving them." In my view, outputs are typical for goal achievement (effectiveness) whereas invested money, skills, workforce, and other resources are more related to the amount of effort (in a broad sense) (Eason, 1988) that is put into the process. In a computer task, such resources could be cognitive capabilities invested in the speed-accuracy trade-off during task execution (Oviatt et al., 2004; Käki, 2004).

Comparing inputs (e.g., effort) to outputs (i.e. the number of completed products) is more of a measure of cost-effectiveness or a return-upon-investment estimation than a description of efficient throughput of materials or information *during process execution*.

It is a bit awkward that effort is not an integral part of the ISO 9241-11 definition of usability because ISO 9241 does have separate sections on the ergonomic requirements of, for instance, visual displays (Part 3), keyboards (Part 4), and work environment (Part 6). Bevan et al. (1991) already stated that usability is largely dependent on the ease of use of a product both ergonomically and as mental effort. Bevan et al. (1991) rely on the ISO/IEC 9126 (1991) standard for software qualities, which says that usability is "a set of attributes of software which bear on the effort needed for use and on the individual assessment of such use ..." They cite Eason (1988) in that ease-of-use is "the degree to which users are able to use the system with the skills, knowledge, stereotypes and experience they can bring to bear."

Effort, then, is not an aspect of effectiveness and efficiency because in itself, effort is goal and process independent. Although effort can be related to, for instance, efficiency (Oviatt et al., 2004), a goal can be achieved without any effort (being lucky) or not achieved despite the effort (bad luck). A process can be efficient although it takes a lot of energy to execute it (e.g., air transport) or a process can be inefficient although the work it takes is not hard (e.g., doing many small work-arounds). Effort is the individual experience of the energy it takes to execute a process so to achieve a certain goal. This effort could be real work power or the symbolic representation of such power, such as money. In many cases, stakeholders want this effort to be as little as possible (easy to use, understandable, comfortable, minor workload, etc.) (Wickens et al., 2004, p. 149).

Particularly in game applications, however, the effort it takes to master the skills or to beat the opponent is highly appreciated (Lauesen, 2005, p. 10). In other words, the effort is not only negative and something to be avoided but is considered optimal (Yerkes & Dodson, 1908) according to a subjective criterion value of the stakeholder. The default level of acceptable effort can be determined by the genre of the application (e.g., game vs. tutorial). In terms of the assembly line, effort is the amount of energy the engine takes to put and keep the assembly belt into gear.

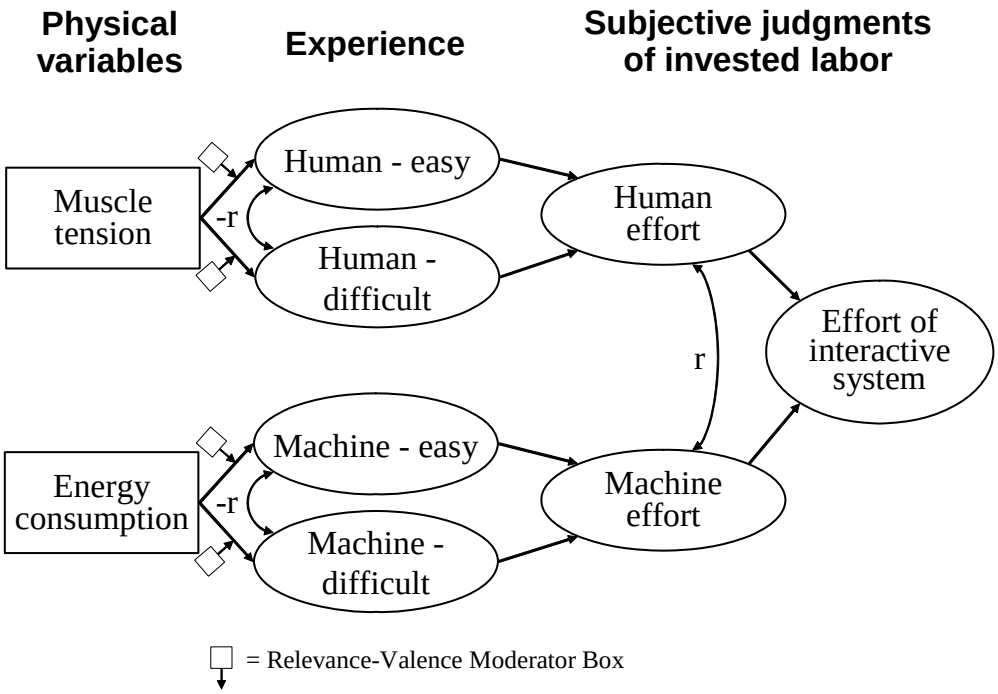


Figure 4: Explanation of judgments on effort expended in working with an interactive system.

In Figure 4, two indicators of human and machine work power are depicted, muscle tension and energy consumption, respectively. Yet, these may not be the only physical variables that could be filled in. Money, number of employees, barrels of oil, Kilojoules, Watt seconds, or other indicators of energy intake or resource depletion may do just as well. Here also, although the physical countable variables lie on a continuum, the experience of such variables is discontinuous – that is, opposite evaluations can occur in parallel. A business transaction may take a certain amount of effort. Yet, if this transaction is important (Relevance is high) and profitable (positive Valence), the effort is seen as little compared to the same effort invested in a transaction that is considered irrelevant and that promises less gains. In other words, if humans or machines have a hard time in executing a process or achieving a goal, the job may be evaluated as difficult, as mixes of difficult and easy, or perhaps even easy, dependent on the levels of Relevance and Valence (Section 5.7). In Figure 4, difficulty and ease stand in a negative relation (-r) to one another, which does not mean that an increase in difficulty is related to the same amount of decrease in ease and v.v.

Human effort and machine effort stand in a positive relation to one another. If the machine has an easy job in task execution, the human will feel that his/her workload is less than if the machine has a tough job. Together, human effort and machine effort determine the overall judgment of Effort invested in an interactive system.

As mentioned, effort has an optimum (inverted-U curve, cf. Yerkes & Dodson, 1908), which need not be minimal. In sum, then, Usability of an interactive system is a curvilinear function of Effort, together with a linear function of Effectiveness, and a linear function of Efficiency (Figure 1).

5.7 Relevance and Valence

The previous chapter reported empirical evidence that Valence has a moderating effect on agreement to requirements on an interactive system. The effects of Relevance were not investigated in that study but for simplicity, I assume that Relevance also acts as a moderator.[4] That is, after the stakeholder has encoded the physically countable variables, these are compared with the three main performance requirements (which pertain to Effectiveness, Efficiency, and Effort) for importance (Relevance) and anticipated gains or losses (Valence). Thus, human and machine should establish a sufficient number of target hits (effectiveness), the right combination of speed and accuracy during process execution (efficiency), at an optimal level of effort. With regard to efficiency, Figure 5 has filled in the speed-accuracy combination with minimal time against minimal error because this is the state that is preferred most of the time.

Features of the human-machine system that are considered relevant to (one of) the three performance requirements evoke more intense reactions (Figure 5, drawn arrows – large effects) than features that are judged irrelevant (Figure 5, dashed arrows – small effects). For example, if connecting to the Internet through broadband triples processing speed, broadband is a relevant feature. An aesthetic GUI – although pleasurable in itself – is an irrelevant feature with respect to the requirement of efficiency. When a connection is established or interrupted due to broadband, broadband evokes more intense responses than a change in the appearance of the dialog boxes during process execution.

System features are also evaluated for their potential to facilitate or inhibit (Valence) the three performance requirements. If an increase in Internet processing speed is most wanted, the outcome expectancy with regard to broadband will be positive and with regard to an analog telephone line, it will be negative. If the attitude towards a feature is positive, that feature contributes (Figure 5) to the level of success (Figure 2), efficiency (Figure 3), or ease (Figure 4) of using the system. If the attitude is negative, the feature contributes (Figure 5) to the level of failure (Figure 2), inefficiency (Figure 3), or difficulty (Figure 4).

[4] A moderating variable influences the level of but is not the cause of, for example, an efficiency score. In Figure 5, the cause of the efficiency score is the observed clock time against error rate. Unlike a moderator, a mediating variable is a necessary step to arrive at a judgment (Baron & Kenny, 1986). In Figure 1, for example, one can get from efficiency to satisfaction only through usability (the mediator).

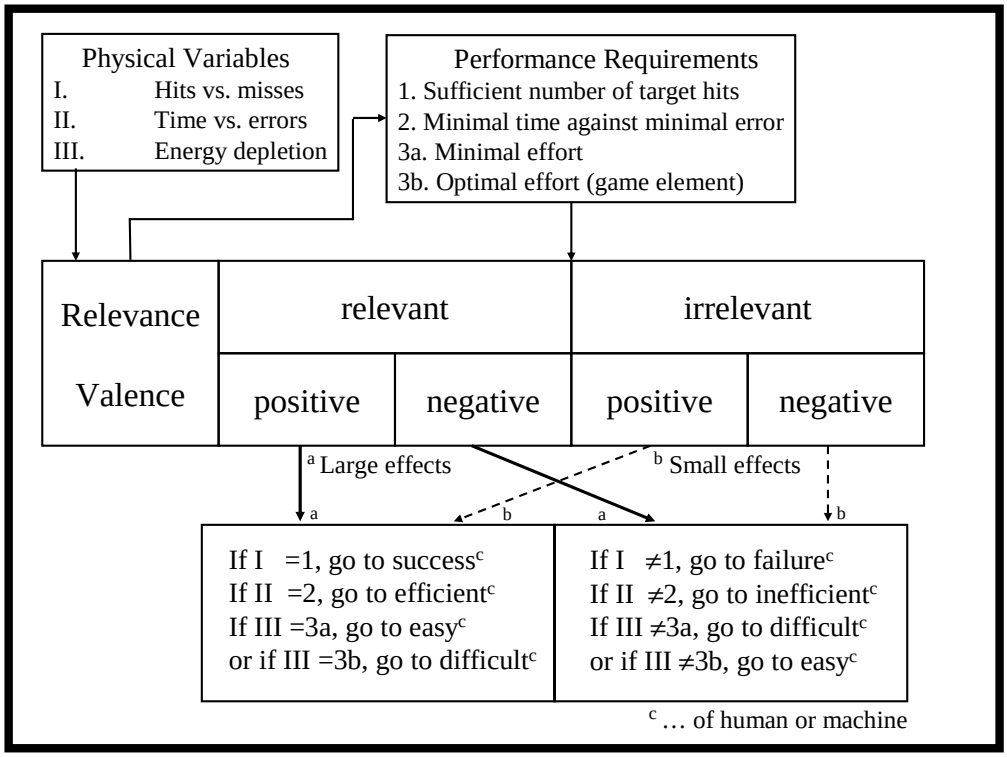


Figure 5: The Relevance-Valence Moderator Box. Physical variables are evaluated on importance and expected gains or losses with regard to the three main performance requirements.

5.8 Conclusions on Theory Development

So far, this chapter has made an attempt to improve the conceptual precision of several important notions in CHI, Human Factors, requirements, and usability engineering. As an important result, models have been developed in which concepts such as performance, usability, and satisfaction have found their place. The triplet effectiveness, efficiency, and effort have been made more precise – from physical foundation to psychological experience – and their mutual relationships have been pointed out.

The model of Stakeholder Logistics (Figure 1) was introduced to better assess the usability of an interactive system in terms of a combination of effectiveness, efficiency, and effort of both humans and machines. The three performance factors presented in Stakeholder Logistics were elaborated in three related sub models. Setting subjective stakeholder criteria to the three performance factors allowed for formulating the three main performance requirements. Stakeholders probably require maximum goal achievement (effectiveness), the highest possible speed against the highest possible accuracy during task or process execution (efficiency), at the cost of an optimal (which could be minimal) level of effort. When tested empirically, averages over stakeholder groups could unveil the critical value of the subjective criteria.

Moreover, the distinction between the human side of the system and the machine aspects has made it possible to understand sometimes-contradictory judgments about the usability of an interactive system. It may well be that a machine is inefficient in executing a task but due to the clever work-arounds of the user, the overall performance of the system may seem to be all right.

The addition of the 'emotion' variables relevance and valence has made it possible to explain that the same quantity of a physical variable can yet evoke different or even contradictory experiences. Dependent on the level of importance and the positive or negative outcome expectancy, the time to achieve a desired goal, for instance, is experienced as longer than the same amount of time that is needed to achieve an undesired goal. Thus, improving a system on the physical level (e.g., faster processing time, more precise algorithms) may not automatically lead to higher levels of experienced efficiency or more satisfaction on the side of the stakeholder.

The models proposed in this chapter have been made open to empirical verification. In the remainder of this chapter, I report on two studies. The first case is a survey of employees who judge a range of banking systems. The second is a structured questionnaire study with anesthetists who evaluate the process efficiency (humans and machines) in the operating room. The (sub) models presented in the theoretical sections each could easily justify four years of Ph.D. research. Therefore, the empirical studies presented next merely test a selection of the assumptions (i.e. on efficiency, usability, and satisfaction) and the rest is regrettably left to future research.

5.9 Study 1: Effectiveness, Efficiency, Usability, and Satisfaction with 25 Banking Systems

5.9.1 *The Usability-Satisfaction Hypotheses*

In Section 5.5, I hypothesized that Satisfaction with an interactive system depends on its Usability, which in its turn depends on Effectiveness, Efficiency, and Effort (Stakeholder Logistics, Figure 1). The latter three variables are assumed to partially correlate. However, authors such as Brooke et al. (1990) and Frøkjær et al. (2000) state the opposite that together with Effectiveness and Efficiency, Satisfaction explains Usability (the ISO 9241-11 definition).

5.9.2 *Method*

5.9.2.1 Participants, Systems, and Procedure

To verify these claims, I was allowed to use the data sampled from four subsequent years (2001-04) of nation-wide user satisfaction research by an ICT department of a multinational bank in The Netherlands (Hoorn, 2005b, Tech. Rep. [CD]).[5] In total, 1943 employees from eight different departments (e.g.,

[5] Brenda Neuteboom of the ICT IDM Management of the bank is kindly thanked for making her data available.

Bank Shops, Advice Offices, Fulfillment, Sales Management, and Credit Support) participated in a yearly conducted electronic query (ITO 2001-04) about 25 interactive systems used in the organization (e.g., Signature Authorization, I-forms/HTML-forms, Card/PIN Activation, Business Transactions, etc.). This ITO survey consisted of 195 single items and open-ended questions of which a subset of 75 items was useful for the present purposes. These 75 items pertained to Satisfaction, Usability, Effectiveness, and Efficiency. Unfortunately, no items on Effort were included in the list.

5.9.2.2 Measurements

Satisfaction was measured by 25 items on interactive banking systems, which were rated on a ten point scale (1= bad, 10= excellent). Not every bank employee used every system that was possibly available. A worker in the bank shop may operate the Mortgage Advice Program but not Peoplesoft Rollout Complaints. To yet obtain an overall measure of Satisfaction with the available interactive systems, I calculated for each employee an average Satisfaction with the system(s) s/he did use (grand mean average Satisfaction, M= 6.49, SD= 1.06).

Usability of the systems was measured by the item "The PC and systems that I work with are sufficiently user friendly to do my work properly," rated on a 6-point scale (1= completely disagree, 6= completely agree) (M= 3.74, SD= 1.20). Effectiveness was supposedly keyed by the item "The PC and systems that I work with provide sufficient information to do my work properly,"[6] also rated on a 6-point scale (M= 4.38, SD= 1.04). Efficiency was supposed to be indicated by the item "The PC and systems that I work with allow me to quickly and adequately accommodate my client's questions and needs," rated for agreement on a 6-point scale (M= 3.67, SD= 1.34).

5.9.3 *Analysis and Results*

5.9.3.1 Testing the ISO 9241-11 Definition of Usability

The ISO 9241-11 norm takes usability as a central concern, which is explained by all other factors (Figure 6). To investigate whether Usability of a set of 25 interactive systems depended on user Satisfaction, Effectiveness, and Efficiency (the ISO definition), I conducted a multiple linear regression analysis. Because not every bank employee used each system, excluding missing data case wise reduced the sample size to N= 928.

Before running the regression analyses, I checked the homogeneity of variance of the variables for the eight different departments. Using the untransformed scores, I calculated the Levene statistic based on the mean, on the median, on the median with adjusted df_2, and on the trimmed mean. However, none of the effects were significant (Levene< 1.25, p >.276) so that variance between departments can be considered homogeneous.

[6] After all, getting information is the main goal of working with an information system.

The regression analysis simulated the ISO definition depicted in Figure 6. That is, Usability was supposedly explained from Satisfaction, Effectiveness, and Efficiency as independent predictors.

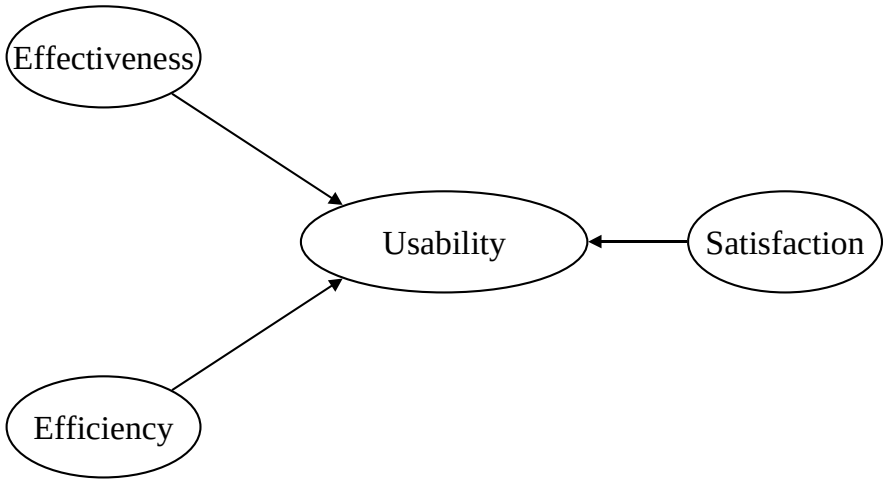


Figure 6: Usability as a central concern, explained by Satisfaction, Effectiveness, and Efficiency, the ISO 9241-11 definition.

In the regression analysis (method Enter), the dependent was Usability of PC and Systems. The controls in step 1 were Information provided by PC and Systems (i.e. Effectiveness) and Working quick and adequately with PC and Systems (i.e. Efficiency). The predictor in step 2 was mean Satisfaction with Interactive Systems.

Efficiency and Effectiveness together accounted for a significant amount of the Usability variability, R^2= .34, R^2_{adj}= .34, $F_{(1,925)}$= 236.58, p= .000. Mean Satisfaction incremented the explained variance of Usability only slightly, R^2_{change}= .04, $F_{change}(1,924)$= 51.62, p= .000. The relative importance of Effectiveness, Efficiency, and Satisfaction in explaining Usability showed that Efficiency was most strongly related to Usability, standardized β= .35, t= 10.15, p= .000. Supporting this conclusion is the height of the standardized Beta coefficient and the strength of the positive correlation between Efficiency and Usability partialling out the effects of all other predictors ($r_{partial}$= .32, r_{part}= .26). Effectiveness was the second best predictor (standardized β= .25, t= 7.96, p= .000, $r_{partial}$= .25, r_{part}= .21). Although close to Effectiveness, mean Satisfaction was the last in line (standardized β= .21, t= 7.18, p= .000, $r_{partial}$= .23, r_{part}= .19).

To make a comparison, I reran the multiple linear regression analysis (method Enter) but this time with Satisfaction as the dependent variable (replacing the central concern of the ISO definition). Thus, the dependent variable was mean Satisfaction with Interactive Systems and the control in step 1 was Usability of PC and Systems. The predictors in step 2 were Information

provided by PC and Systems, indicating Effectiveness, and Working quick and adequately with PC and Systems, indicating Efficiency.

Usability accounted for a reasonable amount of the mean Satisfaction variability, R^2= .16, R^2_{adj}= .16, $F_{(1,926)}$= 181.85, p= .000. Effectiveness and Efficiency incremented the explained variance of Satisfaction only slightly, although this contribution was significant, R^2_{change}= .04, $F_{change}(2,924)$= 25.83, p= .000. With correlation-regression analyses, the relative importance of Usability, Effectiveness, and Efficiency in predicting Satisfaction was assessed. It seemed that Usability was most strongly related to mean Satisfaction, standardized β= .26, t= 7.19, p= .000. Supporting this conclusion is the height of the standardized Beta coefficient and the strength of the positive correlation between Usability and Satisfaction partialling out the effects of all other predictors ($r_{partial}$= .23, r_{part}= .21). Efficiency was the second best predictor (standardized β= .21, t= 5.76, p= .000, $r_{partial}$= .19, r_{part}= .17), whereas Effectiveness offered little or no additional predictive power (p > .05) beyond that contributed by Usability and Efficiency (this finding counters Scott, 1995, but is nevertheless not a confirmation of Pather et al., 2003).

In conclusion, Satisfaction with the banking systems depended more on their Usability (R^2= .16, regression 2) than the other way round (R^2_{change}= .04, regression 1). Yet, the contribution of Satisfaction to Usability was significant, which may indicate that after people are satisfied with a system, their default level of estimated usability of that system is increased when they start using that system for the second or the third time. Put differently, when the process of using the system starts all over again.

In addition, Efficiency and Effectiveness significantly predicted the variance of Usability (R^2= .34), whereas Satisfaction made a minor contribution (see previous paragraph). Efficiency and Effectiveness did so in a relatively independent way (β= .35, $r_{partial}$= .32, r_{part}= .26 for Efficiency and β= .25, $r_{partial}$= .25, r_{part}= .21 for Effectiveness).

In all, these results seem to counter the relations proposed by the ISO 9241-11 definition of usability. Usability was not explained so much by Satisfaction as Satisfaction was by Usability. This leads to testing the Stakeholder Logistics conception of usability.

5.9.3.2 Testing Stakeholder Logistics

To test whether Usability played a mediating role between estimates of Efficiency and Effectiveness on the one hand and experienced Satisfaction on the other (Stakeholder Logistics, Figure 1), I ran a Sobel test for mediation (Sobel, 1982; Preacher & Hayes, 2004). Effectiveness and Efficiency together served as the predictor, Usability as mediator, and Satisfaction as the dependent. As expected by Stakeholder Logistics, the mediation tests revealed indirect effects of Effectiveness and Efficiency on Satisfaction, indicating partial mediation by Usability (Sobel z= 8.03, p< .0001). However, this means that Effectiveness and Efficiency also *directly* contributed to Satisfaction (Figure 7), which was

not expected. The Pearson correlation between Efficiency and Effectiveness was r= .54**. To explore the individual predictive power of Effectiveness and Efficiency, I ran two separate Sobel tests, the results of which are in Figure 7.

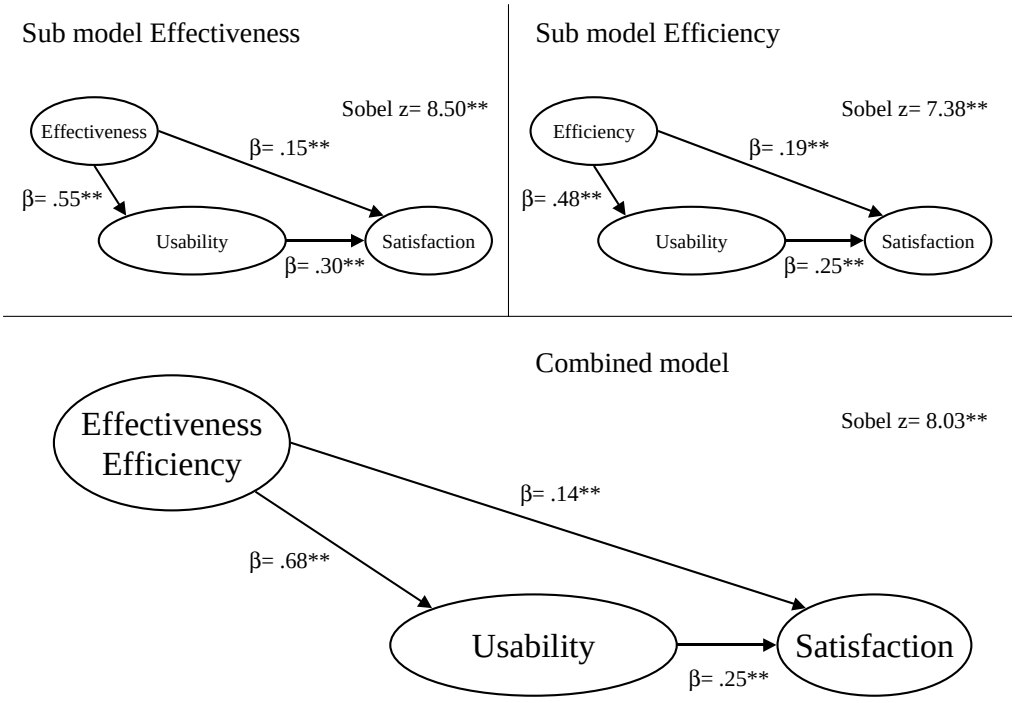


Figure 7: Results of Sobel tests for mediation by Usability (N= 928). All effects significant at $p < .0001$.

5.9.4 *Conclusions for the ISO 9241-11 Definition of Usability versus Stakeholder Logistics*

The constellation proposed by the ISO 9241-11 definition of usability could marginally be retrieved in the data of bank employees judging 25 interactive banking systems over four years. Conversely, Satisfaction was better explained by Usability than the other way round. In addition, Stakeholder Logistics was well retrieved in the data, with a small modification.

Usability indeed was a mediator between the dependent variable Satisfaction on the one hand and the predictors Effectiveness and Efficiency on the other. As predicted, the latter two were relatively independent, which is manifested by a high significant correlation (r= .54) and yet independently maintaining explanatory power. The direct effects of Effectiveness and Efficiency on Satisfaction were not expected by Stakeholder Logistics and all together, the results 'push aside' a little the central concern that ISO 9241-11 attaches to usability.

5.9.5 *The Efficiency Hypothesis*

Next, I will explore some of the assumptions on efficiency put forth in Section 5.6.2 (Figure 3). There I assumed that efficiency judgments are based on experiences of the speed and accuracy of process execution. In an interactive system, such judgments can apply to the human influence on the system and/or to the machine's influence.

5.9.6 *Method*

5.9.6.1 Measurements

To measure Efficiency, in Section 5.9.2.2 I used the item "The PC and systems that I work with allow me to quickly and adequately accommodate my client's questions and needs," rated for agreement on a 6-point scale (M= 3.67, SD= 1.34). To see whether this item could be explained by experiences of speed and accuracy, I selected two more items from the ITO 2001-04 survey list that explicitly mentioned time and error aspects encountered in the task environment. Speed was indicated by "How satisfied are you with the waiting time at the Helpdesk telephone line" (M= 4.42, SD= 1.26). Accuracy supposedly was keyed by "The PC and systems that I work with are so error free that I can do my job properly" (M= 2.86, SD= 1.26).

5.9.7 *Analysis and Results*

5.9.7.1 Testing the Efficiency Hypothesis

In multiple linear regression analysis (method Enter), the dependent variable was the Efficiency item on working quick and adequately with PC and Systems. The predictors were the Speed item about the waiting time at the Helpdesk telephone and the Accuracy item about PC and Systems being error free, which showed a trend for heterogeneous variance according to Levene (based on the mean, p= .051; based on the trimmed mean, p= .074).

The Speed and Accuracy item together accounted for a significant quantity of the Efficiency variability, R^2= .21, R^2_{adj}= .21, $F_{(1,758)}$= 101.12, p= .000. Accuracy was most strongly related to Efficiency, despite the trend for heterogeneous variance, standardized β= .43, t= 13.22, p= .000. Supporting this conclusion is the height of the standardized Beta coefficient and the strength of the correlation between Accuracy and Efficiency partialling out the effect of the other predictor ($r_{partial}$= .43, r_{part}= .43). To a lesser degree, Speed also contributed independently to Efficiency (standardized β= .10, t= 2.94, p= .003, $r_{partial}$= .11, r_{part}= .10). These results support the conclusion that Efficiency indeed is a function of Accuracy and relatively independent of that, a function of Speed.

5.9.7.2 Testing the Speed-Accuracy Hypothesis

In Section 5.6.2, Figure 3, I assumed that experiences of 'fast' and 'accurate' together explain judgments of interactive systems being 'efficient,' whereas

'slow' and 'inaccurate' would feed judgments of systems being 'inefficient.' To arrive at different groups of bank employees that would represent the judgments efficient vs. inefficient, fast vs. slow, and accurate vs. inaccurate, I performed a median split for the Efficiency item (Median= 4.00, Cum%= 68.3), Speed item (Median= 4.00, Cum%= 46.9), and Accuracy item (Median= 3.00, Cum%= 68.4).[7] The employees who judged that the Efficiency was high thus established the score for Efficient; those who thought Efficiency was low established the score for Inefficient. Likewise for the Speed item, which was divided into a score for Fast and a score for Slow as well as the Accuracy item, which was divided into a score for Accurate and a score for Inaccurate.

To investigate whether judgments of efficiency depended on experiences of 'fast' and 'accurate' and inefficiency depended on 'slow' and 'inaccurate,' I performed two multiple linear regressions (method Enter). In analysis (A), the Efficient score served as dependent, the Fast and Accurate scores being the predictors, and Slow and Inaccurate being the controls. In regression analysis (B), the Inefficient score served as dependent, the Slow and Inaccurate score being the predictors, and Fast and Accurate being the controls.

The model assessed with regression (A) was insignificant ($F < 1$). However, the model assessed with regression (B) did hold. Slow and Inaccurate accounted for a reasonable percent of the Inefficient score variability, $R^2= .11$, $R^2_{adj}= .10$, $F_{(2,251)}= 14.82$, $p= .000$. The Inaccurate score was most strongly related to the Inefficient score, standardized $\beta= .28$, $t= 4.61$, $p= .000$, i.e. when partialling out the effect of all other predictors ($r_{partial}= .28$, $r_{part}= .28$). To a lesser extent, Slow also contributed independently to the Inefficient score (standardized $\beta= .14$, $t= 2.28$, $p= .023$, $r_{partial}= .14$, $r_{part}= .14$).

5.9.8 *Conclusions Efficiency Hypothesis*

The data from the selected ITO 2001-04 survey items indeed supported the assumption that Efficiency judgments were based on experiences of the Speed and Accuracy of process execution (Section 5.6.2, Figure 3). Regression analysis showed that the Speed item and Accuracy item significantly accounted for a part of the variability in agreement to the Efficiency item. Moreover, it seemed worthwhile to decompose Speed and Accuracy into experiences of Fast vs. Slow and Accurate vs. Inaccurate, respectively. As expected by the theory (Section 5.6.2), the Slow score and the Inaccurate score significantly contributed to the Inefficient score. Yet, no evidence was obtained for the assumption that the Fast and Accurate score could explain the Efficient score. It might be that the bank employees experienced the systems they worked with merely as inefficient and that considerations of efficiency were absent.

[7] Although certain authors (MacCallum et al., 2002) oppose to the median-split procedure as a loss of information, in this case, there was no alternative to verify hypotheses of unipolarity.

5.9.9 *Discussion Study 1: Interactive Banking Systems*

The assumptions of Stakeholder Logistics (Figure 1) were corroborated and improved by survey data gathered from 928 employees of a multinational bank in The Netherlands. Different from the standard usability definition of ISO 9241-11 (1998) and the findings by Frøkjær et al. (2000), in that banking company, Satisfaction with the interactive systems depended on Usability, whereas Usability depended less on Satisfaction. Moreover, as predicted, Efficiency and Effectiveness explained a significant part of the variance of Usability (regrettably, there were no questions on Effort in this survey). In support of Stakeholder Logistics, Efficiency and Effectiveness were positively correlated (r= .54) but could still independently explain part of the variance of Usability.

Nonetheless, Usability was not the only contributor to Satisfaction and Effectiveness and Efficiency could only partly compensate for the remaining variance to be explained. This shows that there are other satisfiers in a system, perhaps effort, perhaps aesthetics, perhaps task conformance. Yet, the first was not measured whereas the latter two fall outside the scope of Stakeholder Logistics, which most of all is a model of performance evaluation. In this respect, the assumption was corroborated that Efficiency judgments were based on experiences of the Speed and Accuracy of an interactive system. Multiple linear regression analyses showed that Speed and Accuracy significantly explained Efficiency. In addition, the proposed decomposition of Speed and Accuracy into Fast vs. Slow and Accurate vs. Inaccurate was fruitful: Slow and Inaccurate significantly predicted part of the Inefficient score. However, Fast and Accurate could not significantly explain the variability of the Efficiency score, which can perhaps be explained as a bias of the stakeholders to the experienced inefficiency of the systems.

From these results, the following picture emerges (Figure 8). Usability explained Satisfaction more than the other way round. Moreover, Usability depended on at least Effectiveness and Efficiency, the latter two being partially correlated. In addition, Effectiveness and Efficiency also directly influenced Satisfaction. Further, Efficiency of the bank's IT depended on measures of Speed and Accuracy, as predicted. Stakeholders that deemed the IT Inefficient thought so because the processes were regarded as Slow and above all, Inaccurate. However, no significant evidence was found that IT that was considered Efficient depended on experiences of processes being Fast and Accurate (gray dashed).

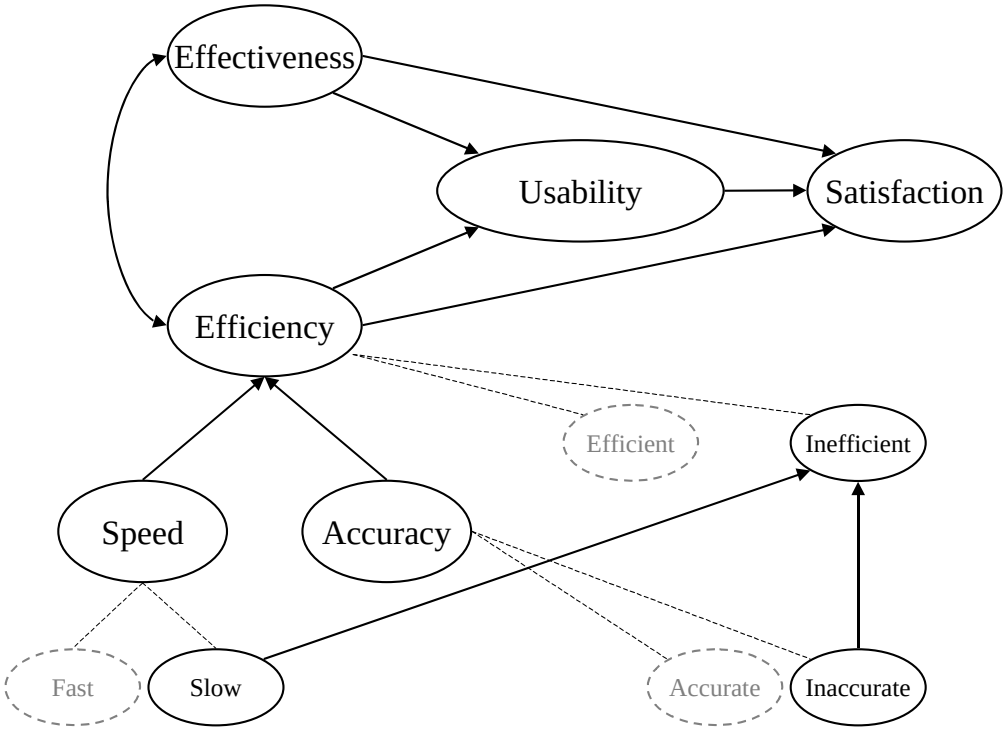


Figure 8: Stakeholder Logistics of 928 users of 25 interactive banking systems.

However, the data used for hypothesis testing were part of the organization's internal evaluation of the IT and IT use. They were not gathered for theoretical purposes. Moreover, the data were sampled with single items, which makes it impossible to do reliability analysis of the measurements. In other words, the (absence of) empirical evidence in this data set should be considered as preliminary. To gather more reliable data, Hoorn and Klaasse (2006, Tech. Rep. [CD]) constructed a structured questionnaire to investigate the business processes in the operation room (OR) of an academic medical center. This questionnaire was used to set up the performance requirements of an electronic efficiency monitor, consisting of an input module and an electronic score board. With it, the throughput of patients and OR allocation time was supposed to be optimized.

5.10 Study 2: Efficiency, Usability, and Satisfaction in the Operating Room

Medical centers (MCs) in The Netherlands have problems with reducing waiting lists. One solution that the academic MC we worked with opted for was to improve the allocation of available space (i.e. the operation rooms - ORs) and the time spent in there. The academic MC has around 700 hospital beds and over 35,000 intakes a year. It is the workplace of about 5000 employees and has a budget of 350 million Euros a year. The operation complex is distributed

over two floors and in total there are about 18 ORs available. A problem in this MC was that the OR planners and the nurses at the ward did not or insufficiently knew how far an operation had progressed. Therefore, they could not optimally anticipate to changes in the schedule, operations running late, ending early, or that were canceled after all. Too often, ORs remained empty where they could have served another patient.

Therefore, the MC was in the process to develop an electronic "score board" that would be connected to an input module at the anesthesiologist's device to inform the nurses at the ward and the OR planners about the operation's progress. Together with my master student Bob Klaasse, I stepped in to set up the performance requirements for this information system, with a particular focus on matters of time and place, that is, the optimization of the OR's efficiency.

For theoretical purposes, however, we did not only perform the requirements engineering but I wanted to know in how far the stakeholders of this progress monitoring-system inside the OR (i.e. the anesthetists and anesthesiologists) *experienced* the operation process as efficient or not. This information is important because it may be possible to reduce the operation process-time physically but if the stakeholders think they already are at the limit of what is possible, making the process faster may evoke quite some dissatisfaction and may eventually lead to rejecting the system.

5.10.1 *Method*

We studied the experience of an operation's efficiency by running a structured questionnaire with a sample of anesthetists in the ORs of an academic MC (Appendix 5.1). Yet, we first wanted to establish a context of use and examine the stakeholders' tasks and goals (ISO 9241-11; Jokela et al., 2003) by performing in-depth ethnography. Bob Klaasse was even allowed to attend an open-heart surgery.

5.10.1.1 Participants

Table 1 gives an overview of all the staff involved in performing an operation. However, our main concern was with the anesthetists and anesthesiologists because they would provide the input to the score board.

Commonly there is one anesthesiologist responsible for the whole operation. When s/he thinks the patient cannot handle the operation, it is cancelled. This anesthesiologist is usually responsible for multiple operations at the same time and will therefore not be present at the whole operation. In stead, an assistant anesthetist monitors the patient while the anesthesiologist occasionally drops by.

The assistant anesthetist and the anesthesiologist are responsible for the comfort and safety of the patient. Shortly before the operation the patient must feel at ease, eventually with help of pre-medication. Then the patient gets the main medication to start the actual anesthesia. During the operation the pa-

tient's vital functions are monitored (e.g., hart beat, blood pressure, breathing, temperature, liquid value, oxygen value in the blood). After the operation, the patient recovers from the anesthesia and should take up breathing again. The anesthetist guides the patient through this process (Heitmiller, 2005). We focused on the anesthetists that were present during the whole operation, because they would be the ones to input the data on the operation's progress and so their judgments on efficiency would be most informative. The 32 participants who filled out our questionnaire consisted of 11 anesthesiologists in training, 10 assistant anesthetists, and 11 anesthetists.

Table 1: Staff active in the operation rooms.

Team	Function
Surgical team	Chief surgeon Assistant surgeon Scrub nurse
Team of anesthetists	Anesthesiologist Anesthetist assistant
Supporting staff	Circulating nurse X-ray specialist

5.10.1.2 Business Process and Systems

The new operation-progress monitoring-system (input module and score board) should be modeled after the operation process. This is a 'business' process that is strongly affected by the interaction between humans and machines. To provide the score board and input module with the right performance (i.e. efficiency) requirements, we not only investigated the speed and accuracy judgments about people but also about the most commonly used machines in the OR. For a more detailed description of each machine, see Appendix 5.2.

The business process in an OR involves five people at least, who are working simultaneously on different parts of the process. The phases in this process consist of clearly described tasks. Table 2 shows which people and which machines are active for each task and learns that the team of anesthetists is involved in more than 50% of the tasks. Therefore, they have the best view on the operation's progress. The most time-consuming part is the operation itself (Task 12). Expert interviews revealed that the greatest error source is human, particularly while making the X-ray photographs (Task 10). The interaction with the vital signs monitor is problematic. It is regularly connected improperly (human error) and sometimes gives the heart pulse in double time (machine error). Most of the time, this is immediately recognized by the anesthetists, who should re-attach the sensor quickly. There are more general human faults that are non-specific to a task, such as having to refill the syringe, instruments missing, or dropping something on the floor.

Table 2: Time and error aspects of the operation process in relation to human-machine interaction.

Tasks	Humans	Machines	Time (min) [a]	Human Error	Human Error Frequency [c]	Machine Error
1. Inspect devices	Anesthetists team	Anesthesiology device	2	-	-	Input module (stylus and touch screen) hamper Automatic data recording fails due to network error
2. Bring patient to OR	Team of anesthetists	-	2	Wrong OR	Seldom	-
3. Move patient	Team of anesthetists and surgical team	-	1	Patient falls	Seldom	-
4. Attach monitoring	Team of anesthetists	Vital signs monitor	3	Improperly connected	Regularly	Gives heart pulse in double time
5. Start introduction	Team of anesthetists	Respiratory ventilator, vaporizer or infusion pump	3	Wrong dosage of anesthesia	Regularly	-
6. Position patient	Team of anesthetists and surgical team	-	3	-	-	-
7. Wash hands	Surgical team	-	6	Touching something non-sterile	Seldom	-
8. Cover up patient	Surgical team	-	3	Improper coverage	Seldom	-
9. Prepare X-ray device	X-ray specialist	X-ray device	5	Forgotten something	Seldom	-
10. Make X-ray photographs	X-ray specialist	X-ray device	7	Low quality of photo	Often	-
11. Begin incision	Surgical team	Suction device Coagulator	5	-	-	-
12. Perform operation	Surgical team	Suction device Coagulator	38 [b]	Unexpected trouble	Regularly	-

13. Close wound	Surgical team	Suction device	10	Improper closure	Seldom	-
14. Start outroduc-tion	Team of anesthetists and surgical team	Vaporizer or infusion pump	5	Patient does not wake up at once	Regularly	-
15. Move patient	Team of anesthetists	-	1	Patient falls	Seldom	-
16. Bring patient to recovery	Team of anesthetists	-	1	-	-	-

[a] Estimates based on expert interviews and informal timing.

[b] Mean duration of 1162 caesarean operations; between 2002 and 2005, the most commonly performed operation at the MC.

[c] Estimates based on expert interviews only.

5.10.1.3 Procedure

Based on our ethnography, we created a structured questionnaire of 61 items. Sixty items divided over seven blocks systematically covered the concepts of the Efficiency model depicted in Figure 3, Section 5.6.2. One item pertained to the respondents' function in the OR. Table 3 shows the order of item blocks in the questionnaire and their contents. The order within and between the Efficiency blocks and the Satisfaction blocks was reversed such that eight different versions of the questionnaire were created. Items were randomized within blocks. In two sessions, the head of the anesthesiologist team distributed the pen-and-paper questionnaires among the anesthetists during their break. Filling out the questionnaire took between 15 and 20 minutes. The questionnaire was in Dutch (Appendix 5.1). The contents of the items are discussed next.

Table 3: Order of item blocks.

Block	Items
1	Introduction letter and question about function
2	Requirements items
3	Usability items
4	1st Speed / Accuracy items
5	1st Efficiency items / Satisfaction items
6	2nd Speed items / Accuracy items
7	2nd Efficiency items / Satisfaction items
8	3rd Speed items / Accuracy items

5.10.1.4 Measurements

The 60 Likert type items systematically covered the theoretical notions by phrasing them positively as well as negatively (Table 4). That is, the number of requirements that should be on the system (Must haves) was counterbalanced by the same number of requirements that should not be on the system (Won't haves). Similarly, satisfaction items counterbalanced dissatisfaction

items, etc. To avoid answering biases, all items were stated affirmatively, avoiding linguistic negations. The head of the anesthesiologist team checked the items for readability and sensibility for someone working in the field. In the main test, anesthetists scored the items for agreement on a 6-point rating scale (0= completely disagree, 1= disagree, 2= disagree a little, 3= agree a little, 4= agree, 5= completely agree).

Table 4: Structure and contents of questionnaire items, following the Efficiency model (Figure 3, Section 5.6.2).

Theoretical Notion	Scales (bipolar)	Sub Scales (unipolar)	Sample Items	#
Accuracy	Human Accuracy	Human Accurate	The instruments are placed in the right way	4
		Human Inaccurate	The instruments are placed wrongly	4
	Machine Accuracy	Machine Accurate	Self-test of anesthesiology device goes correctly	4
		Machine Inaccurate	Self-test of anesthesiology device goes incorrectly	4
Speed	Human Speed	Human Fast	Checking the anesthesiology device goes quickly	4
		Human Slow	Checking the anesthesiology device is time-consuming	4
	Machine Speed	Machine Fast	The respiratory ventilator is readily available	4
		Machine Slow	The respiratory ventilator takes a long time to start	4
Efficiency	Human Efficiency	Human Efficient	OR personnel is well-organized	1
		Human Inefficient	OR personnel is badly organized	1
	Machine Efficiency	Machine Efficient	Machine support runs smoothly	1
		Machine Inefficient	Machine support runs poorly	1
	Overall Efficiency	Overall Efficiency	The session proceeds in good order	1
		Overall Inefficiency	The session proceeds disorderly	1
Satisfaction	Human Satisfaction	Satisfied with Humans	OR personnel works all right	1
		Dissatisfied with Humans	OR personnel works inadequately	1
	Machine Satisfaction	Satisfied with Machines	Machine support is Okay	1
		Dissatisfied with Machines	Machine support is weak	1
	Session Satisfaction	Satisfied with Session	Pleased with progress session	1
		Dissatisfied with Session	Disappointed about progress session	1
Usability	-	-	The coagulator is a handy device	8
Requirements	Must haves Won't haves	- -	Input that the session runs late Warning beeps	4 4
Function in OR	-	-	-	1

5.10.2 *Analysis and Results*

5.10.2.1 Scale Analysis

To explore whether we could trust what we measured, we used Standardized Cronbachs's alpha to establish the average correlation of items within a scale (internal consistency reliability). As a further verification, Corrected Item-Total Correlations showed the correlations of a single item with the sum of all other items. To check whether items were unambiguous or belonged to more scales, we conducted factor analysis because items were not supposed to correlate strongly with other scales.

Item selection was a trade-off among several criteria. For each item, outliers were considered missing values and replaced by the mean of the remaining values. For two Usability items, more than 30% missing values were observed and these items were not used in the analyses. The aim was to establish as many items on a (sub) scale as possible, provided that Standardized Cronbach's alpha for a scale was at least > .60, preferably > .70, that Corrected Item-Total Correlations of each item was > .20, and that items showed the lowest correlations possible with other scales. The Standard Deviation (SD) of items and scales should be around 1. The skewness of items and scales should be < .70 and if present, so-called leverage points were removed.[8] Table 5 shows the results of the thus revised (sub) scales. For the complete scale analysis, consult (Hoorn & Klaasse, 2006, Tech. Rep. [CD]). Table 5 shows that the reliability of Human Speed was insufficient (.42) and further analysis will be conducted with the two single items. The reliability of all other scales that could be established ranged between .64 and .87, which is sufficient to good.

In view of the models of Stakeholder Logistics (Figure 1) and Efficiency (Figure 3), scale analysis (i.e. reliability in combination with factor analysis) reduced the number of possible hypotheses that could be tested (Figure 10). Some of the major deviations from the earlier conception were that Satisfaction was split into a human and a machine component, and that the speed-accuracy decomposition for machines was simplified. For the human side of the process, this decomposition (fast vs. slow – accurate vs. inaccurate) remained intact.

[8] Leverage points are values extremely distant from the center of the sampled predictor values.

Table 5: Standardized Cronbach's alpha, Means and SDs, and the length of the shortened scales (N= 32).

Scale	Standardized Cronbach's Alpha	Mean	Standard Deviation	Items
Must Requirements	.78	4.38	.53	Must 3 – Order patient by computer (skewed -.73) Must 4 – Indicate via computer when pre-medication is due
Won't Requirements	single item	2.09	1.59	Won't 4 – Administer markers on more computers
Satisfaction with Humans	.87	2.56	1.02	Satisfied 1 – with Session Satisfied 2 – with Personnel Dissatisfied 4 – with Session (reverse scaled) (skewed .89) Dissatisfied 5 – with Personnel (reverse scaled)
Dissatisfaction with Machines	.66	3.18	.80	Satisfied 3 – with Machines Dissatisfied 6 – with Machines (reverse scaled)
Usability	.74	4.29	.56	Usable 3 – Infusion pump Usable 4 – Respiratory ventilator
Human Efficiency	.82	2.45	1.17	Efficient 1 – Session Efficient 2 – Personnel
Machine Efficiency	.78	2.81	1.08	Efficient 3 – Machines Inefficient 6 – Machines (reverse scaled)
Session Inefficiency	single item	3.46	.78	Inefficient 4 – Session
Human Accuracy	single item	3.50	.62	Human Accurate 4 – Instruments are placed in the right way
Human Inaccuracy	single item	2.50	.91	Human Inaccurate 8 – People drop things
Machine Accuracy	void	-	-	-
Machine Inaccuracy	.81	2.13	1.26	Machine Inaccurate 7 – Data input goes badly Machine Inaccurate 8 – Automatic registration of anesthesiology data fails
Human Speed	.42	2.70	1.20	Human Fast 4 – Patient is covered quickly Human Slow 7 – Closing the operation wound is a lengthy procedure (reverse scaled)
Machine Speed	.64	2.39	.92	Machine Fast 2 – The coagulator burns the veins quickly Machine Slow 7 – The input module works with delays (reverse scaled) Machine Slow 8 – Start up of the input module takes a while (reverse scaled)

Note: Skewness of the final scales was between -.50 and .50

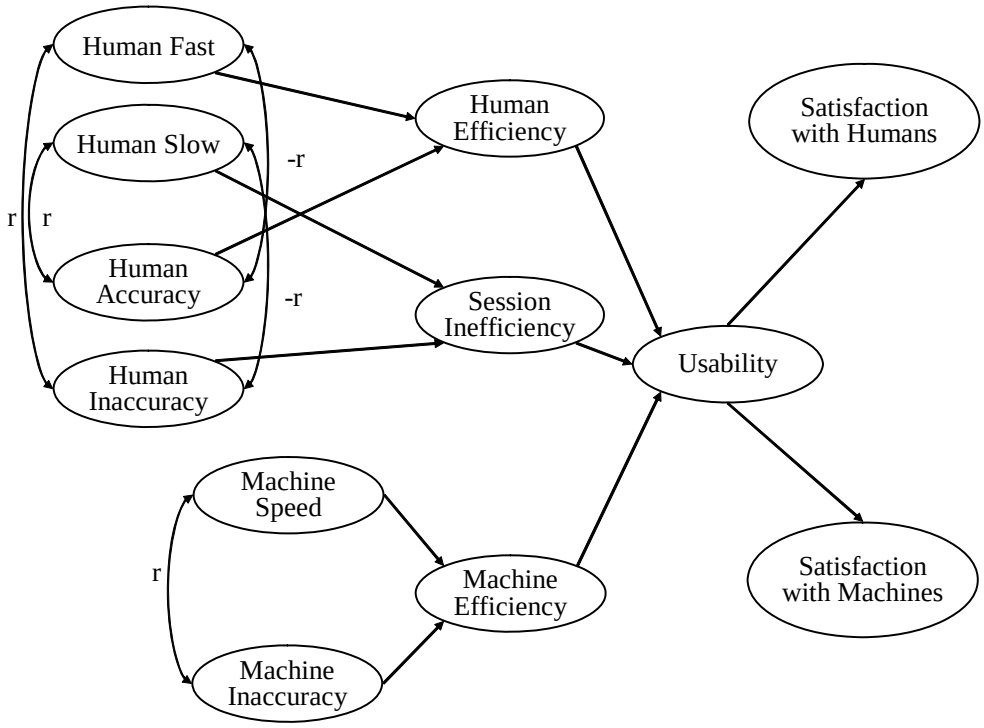


Figure 10: Stakeholder Logistics in relation to Efficiency as far as the data allowed to test.

5.10.2.2 Hypothesis Testing

In following ISO 9241-11 (1998), authors such as Brooke et al. (1990) and Frøkjær et al. (2000) conjectured that satisfaction explains usability whereas Stakeholder Logistics claims the reverse. Next, bullets indicate the hypotheses that were tested by means of (multiple) linear regression analyses (method Enter).

- (Satisfaction with Humans, Satisfaction with Machines) → Usability

In a multiple linear regression (method Enter), Usability served as dependent and Satisfaction with Humans as well as Satisfaction with Machines served as the two predictors. However, no significant effects occurred, R^2= .08, R_{adj}^2= .02, $F_{(2,29)}$= 1.23, p= .307 (for details, consult Hoorn & Klaasse, 2006, Tech. Rep. [CD]).

- Usability → Satisfaction with Humans

Usability served as the sole predictor of Satisfaction with Humans, but no significant effect occurred (F< 1).

- Usability → Satisfaction with Machines

Usability served as the sole predictor of Satisfaction with Machines, but again no significant effect occurred, R^2= .08, R_{adj}^2= .04, $F_{(1,30)}$= 2.42, p= .131.

Thus, Satisfaction with Humans or Machines did not explain Usability, but Usability did not explain Satisfaction with Humans or with Machines either. In this respect, the hypotheses of both ISO 9241-11 and Stakeholder Logistics were rejected. Therefore, I explored whether other variables could predict Usability and this additional hypothesis is tested next.

- (Human Efficiency, Machine Efficiency, Session Inefficiency) → Usability

Machine Efficiency explained a significant amount of the Usability variability, R^2= .28, R_{adj}^2= .20, $F_{(3,28)}$= 3.63, p= .025. To assess the individual strength of the predictors, correlation-regression analyses were performed. Machine Efficiency was the only significant predictor of Usability, standardized β= .55, $r_{partial}$= .53, r_{part}= .53. Human Efficiency and Session Inefficiency did not significantly contribute to the Usability variability.

Usability seemed to be affected by efficiency measures. Therefore, I tested whether the same predictors were influential for Satisfaction.

- (Human Efficiency, Machine Efficiency, Session Inefficiency) → Satisfaction with Humans

Human Efficiency was the only significant predictor of Satisfaction with Humans, R^2= .49, R_{adj}^2= .44, $F_{(3,28)}$= 9.05, p= .000, standardized β= .57 $r_{partial}$= .59, r_{part}= .51. Machine Efficiency and Session Inefficiency did not significantly contribute to the variability in Satisfaction with Humans. Session Inefficiency established merely a trend (p= .07).

- (Human Efficiency, Machine Efficiency, Session Inefficiency) → Satisfaction with Machines

Machine Efficiency was the only significant predictor of Satisfaction with Machines, R^2= .33, R_{adj}^2= .25, $F_{(3,28)}$= 4.52, p= .010, standardized β= .59, $r_{partial}$= .57, r_{part}= .57. Human Efficiency and Session Inefficiency did not significantly contribute to the variability in Satisfaction with Machines.

Instead of Usability, Human Efficiency indeed predicted Satisfaction with Humans. In addition, Machine Efficiency indeed predicted Satisfaction with Machines. Next, I tested the speed-accuracy trade-off expected by the Efficiency model (Figure 3) as far as scale analysis (Table 5, Figure 10) permitted.

- (Machine Speed, Machine Inaccuracy) → Machine Efficiency

Machine Inaccuracy was the only significant predictor of Machine Efficiency, R^2= .41, R_{adj}^2= .37, $F_{(2,29)}$= 9.94, p= .001, standardized β= -.49, $r_{partial}$= -.60, r_{part}= -.49. Machine Speed did not significantly contribute to the variability in the Machine Efficiency measure.

- (Human Fast, Human Slow, Human Accuracy, Human Inaccuracy) → Human Efficiency

Human Inaccuracy and Human Fast explained Human Efficiency, R^2= .33, R_{adj}^2= .23, $F_{(4,27)}$= 3.33, p= .024. Correlation-regression analyses revealed that Human Inaccuracy was the main predictor, standardized β= .44, $r_{partial}$= .46, r_{part}= .42. The second best predictor of the Human Efficiency measure was Human Fast, standardized β= .36, $r_{partial}$= .39, r_{part}= .34. Human Inaccuracy and Human Fast had a negative correlation (r= -.14). Human Accuracy and Human Slow did not significantly explain the variability in Human Efficiency.

- (Human Fast, Human Slow, Human Accurate, Human Inaccurate) → Session Inefficiency

The effects on the Session Inefficiency measure were insignificant (F< 1). These results suggest that inaccuracy measures best predicted the estimated efficiency. Machine Inaccuracy was the sole predictor of Machine Efficiency and Human Inaccuracy predicted Human Efficiency best, followed by Human Fast.

As suggested in Section 5.1, dissatisfaction with the usability of the current system could easily lead to the demand of novel or changed usability features, thus stating a list of new (i.e. usability) requirements. Therefore, I verified whether Satisfaction and Usability affected the level to which stakeholders agreed with the Must and Won't Requirements on the OR-process monitoring-system that was under development.

- (Satisfaction with Humans, Satisfaction with Machines) → Must Requirements

The effects of Satisfaction on agreement to the Must Requirements were insignificant, R^2= .12, R_{adj}^2= .06, $F_{(2,29)}$= 1.99, p= .154.

- (Satisfaction with Humans, Satisfaction with Machines) → Won't Requirements

The effects of Satisfaction on (dis)agreement to the Won't Requirements were insignificant (F< 1). Satisfaction (whether with Humans or Machines) had no explanatory power for Requirements what so ever. Therefore, I turned to all other variables as predictors of agreement to requirements, including Usability.

- (Machine Speed, Human Inaccuracy, Human Accuracy, Session Inefficiency, Human Fast, Human Slow, Usability, Machine Inaccuracy, Human Efficiency, Machine Efficiency) → Must Requirements

All effects on agreement to the Must Requirements were insignificant, R^2= .38, R_{adj}^2= .08, $F_{(10,21)}$= 1.27, p= .307. However, in the correlation-regression analyses, two trends in explaining the Must Requirements variability occurred for Machine Inaccuracy (p= . 06) and Machine Speed (p= .085). This line will be pursued next. Most notably, Usability lacked any explanatory power for Must Requirements.

- (Machine Speed, Machine Inaccuracy) → Must Requirements

Machine Inaccuracy and Machine Speed both accounted for a significant amount of the variability in agreement to Must Requirements, R^2= .22, R_{adj}^2= .17, $F_{(2,29)}$= 4.11, p= .027. Correlation-regression analyses revealed that Machine Inaccuracy was the better predictor, standardized β= .50, $r_{partial}$= .44, r_{part}= .44 followed by Machine Speed, standardized β= .42, $r_{partial}$= .38, r_{part}= .36. The correlation between Machine Inaccurate and Machine Speed was negative (r= -.49**).

- (Machine Speed, Human Inaccuracy, Human Accuracy, Session Inefficiency, Human Fast, Human Slow, Usability, Machine Inaccuracy, Human Efficiency, Machine Efficiency) → Won't Requirements

Again, all effects on agreement to the Won't Requirements were insignificant (F< 1). Yet again, correlation-regression analysis revealed a trend for Human Accuracy to explain Won't Requirements (p= .018). In this case also, Usability could not significantly explain (dis)agreement to Won't Requirements.

- Human Accuracy → Won't Requirements

Human Accuracy significantly accounted for the variability in (dis)agreement to the Won't Requirements, R^2= .13, R_{adj}^2= .10, $F_{(1,30)}$= 4.41, p= .044, standardized β= -.36, $r_{partial}$= -.36, r_{part}= -.36. The relation between Human Accuracy and Won't Requirements was negative. Thus, the more accurate the anesthetists believed they did their work, the less they agreed with the single-item Won't Requirement of "Administering markers on more computers."

Figure 11 provides a summary of the regression results. Three sub models transpired. Surprisingly, judgments about humans and machines did not seem to affect one another. The third sub model truly covered a user-centered design (UCD) approach in that aspects of the system (Won't Requirements) were predicted from measures of Human Accuracy.

Figure 11 (dashed arrows) also shows possible relations that should be studied with mediation tests (Sobel, 1982; Preacher & Hayes, 2004). Table 6 shows the results of the mediation tests, confirming the patterns found with the regression analyses and indicating some trends for direct effects (partial mediation) of Machine Inaccuracy on Satisfaction with Machines and on Must Requirements. In the OR case, mediation was not substantiated significantly for Stakeholder Logistics.

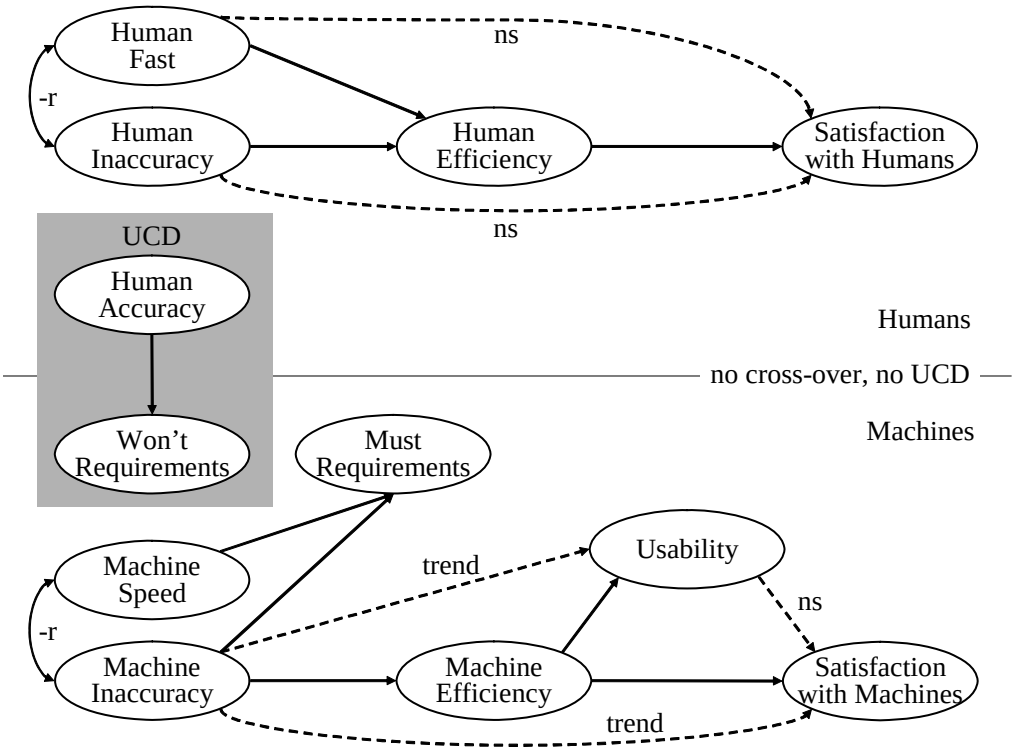


Figure 11. Stakeholder Logistics in the OR based on regression results (drawn arrows). Three sub models emerged, one for Humans, one for Machines, and one for truly user-centered design (UCD). Dashed arrows indicate relations verified with mediation tests.

Table 6: Results of Sobel tests for mediation (N= 32). Trends (p≈ .05) of partial mediation were found for Machine Inaccuracy.

Path Predictor→Mediator→Dependent	**Sobel z**	**p**	**Effect**	**β**	**p**
Human Fast → Human Efficiency → Satisfaction with Humans	1.84	.066	Human Fast → Human Efficiency Human Efficiency → Satisfaction with Humans (controlled for Human Fast)	.33 .59	.0481 .0001
Human Inaccuracy → Human Efficiency → Satisfaction with Humans	1.78	.075	Human Inaccuracy → Human Efficiency Human Efficiency → Satisfaction with Humans (controlled for Human Inaccuracy)	.44 .58	.0156 .0001

Machine Inaccuracy → Machine Efficiency → Usability	-2.01	.045	Machine Inaccuracy → Usability Machine Inaccuracy → Machine Efficiency Machine Efficiency → Usability (controlled for Machine Inaccuracy)	-.15 -.52 .24	.0559 .0003 .0253
Machine Inaccuracy → Machine Efficiency → Satisfaction with Machines	-2.41	.016	Machine Inaccuracy → Satisfaction with Machines Machine Inaccuracy → Machine Efficiency Machine Efficiency → Satisfaction with Machines (controlled for Machine Inaccuracy)	-.20 -.52 .43	.0802 .0003 .0049
Machine Efficiency → Usability → Satisfaction with Machines	-.08	.933	Machine Efficiency → Satisfaction with Machines Machine Efficiency → Usability Machine Efficiency → Satisfaction with Machines (controlled for Usability)	.11 .08 .13	.0008 .0031 .0033

5.10.3 *Conclusions/Discussion Study 2: Operation Room Systems*

The usability-satisfaction hypotheses of both ISO 9241-11 (e.g., Brooke et al., 1990; Frøkjær et al., 2000) and Stakeholder Logistics were rejected. Satisfaction (either with Humans or Machines) did not significantly contribute to Usability but the reverse did not hold either. Human Efficiency, however, predicted Satisfaction with Humans while Machine Efficiency predicted Satisfaction with Machines. Improving performance factors (here, efficiency) apparently is more important than usability to increase stakeholder satisfaction.

The predicted speed-accuracy trade-off (Figure 3) could be retrieved for three instances. (1) Machine Inaccuracy and not Machine Speed was the only significant predictor of Machine Efficiency. This suggests that in the medical context of use, avoiding machine error was more important than increasing the process speed of the machines. The correlation between Machine Inaccuracy and Machine Efficiency was negative (standardized β= -.49), indicating that the anesthetists believed that a decrease in inaccuracy enhanced the efficiency of the machine. Taken together, interpreting efficiency as optimization of time

or speed alone is insufficient (cf. Käki, 2004). For specialists, the absence of machine error seems to have significantly more value than aspects of speed to estimate how efficient their machines are.

(2) Human Inaccuracy and Human Fast explained Human Efficiency. Similar to machine efficiency, human efficiency in the OR was better predicted from the absence of mistakes than from a high work pace. The correlation between Human Inaccuracy and Human Fast was negative (r= -.14), indicating that the anesthetists believed that they became faster when they made fewer mistakes. However, the standardized Betas of Human Inaccuracy and Human Fast were both positive. Thus, speeding up the work pace increased efficiency but the anesthetists realized that this was at the cost of making more errors. In other words, the speed-accuracy trade-off usually made visible in reaction-time paradigms (e.g., Meyer et al., 1988) was also retrievable in subjective judgments on human efficiency. On the one hand, the anesthetists knew that they could gain speed by making less mistakes but they also saw that there was an optimum to this relation, so that capitalizing on speed alone would lead to higher error rates. From a user-centered design perspective, it is a bit weird that the anesthetists felt that human and machine accuracy were unconnected (Figure 11). Table 2, for instance, shows that the input module suffers from flaws that need to be restored by hand. It might be that the anesthetists are so used to different equipment with non-conforming user interfaces that they did not see that they were making up for the flaws of the machine and in this line, that usability was much of an issue.

(3) Machine Inaccuracy and Machine Speed both accounted for a significant amount of the variability in agreement to Must Requirements. Similar to the findings for Machine Efficiency, what must be on the new OR-process monitoring-system was better predicted from the wish to reduce errors in machines than from wanting faster machines. Similar to Human Efficiency, the correlation between Machine Inaccuracy and Machine Speed was negative (r= -.49**), indicating that the anesthetists believed that machines became faster when they were more error-free. However, the standardized Betas of Machine Inaccuracy and Machine Fast were, again, both positive. Thus, when the anesthetists saw more speed in their machines, they agreed more with what the new monitoring system must have. However, they also realized that this was at the cost of getting a machine that was more error-prone. Thus, the speed-accuracy trade-off reflected in the scores on the scales Machine Inaccuracy and Machine Speed accounted for the variability in agreement to the Must Requirements.

What probably happened, then, was this. When the anesthetists saw more inaccuracy in their machines, they agreed more to what must be on the new system – probably to avoid such inaccuracy. When they saw more speed in their machines, they also agreed more to the must requirements – probably to maintain or achieve high speeds. The conflict between these stakeholder goals had sufficient power to explain the changeability in agreement to the require-

ments (should I agree or disagree?). On the one hand, the new system must secure speed but on the other hand, avoid becoming flawed. The fact that a negative goal state (system errors) should be avoided by a positive requirement (Must) is typical for the goals-to-requirements chiasm (Chapter 4). This negative-positive constellation predicts the changeability in agreement to requirements better than straightforward positive-positive relations, such as achieving the desired goal state of increased speed by demanding a positive requirement (Must). Indeed, in view of the strengths of the standardized Betas and part and partial correlations, Machine Inaccuracy (negative) had more predictive power for the variability in agreement to Must Requirements (positive) than Machine Speed (positive). With it, the speed-accuracy trade-off reflected in the expert judgments supplied the goals-to-requirements chiasm with another concrete example. It also became clear, again (cf. Chapter 4, i.e. LWMS), that lower-level goals such as enhanced system performance rather than higher-order goals such as usability feed requirements change.

In addition, Human Accuracy accounted for the Won't Requirements. Again, this was a manifestation of the goals-to-requirements chiasm (Chapter 4). To preserve accuracy (which probably was the positive goal to achieve), the negative requirement of data input on multiple computers was rejected. The negative relation between Human Accuracy and Won't Requirements was the only place where considerations about humans nurtured judgments about machines. In all other cases, the thoughts about humans ran parallel to but never connected with those about machines (Figure 11). It seems, then, that for these stakeholders, user-centered design was not an issue but efficiency was – of humans and independently of that, machines.

Using quite a different type of questionnaire construction than in Chapter 4, the goals-to-requirements chiasm could again be established (Figure 12). In all, it seems that usability played a subordinate role in the 'User Virtual Machine' of these stakeholders. In other words, requirements engineers should not focus on usability requirements but rather on *performance requirements of the machine!* Usability was a non-issue because stakeholders cared more about their personal and machine logistics.

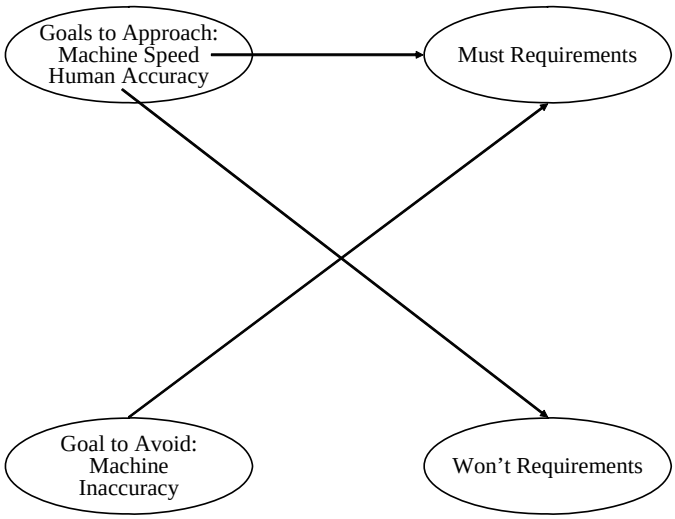


Figure 12: Goals-to-requirements chiasm found in the OR case.

5.11 General Discussion

In dealing with several key concepts in CHI, interaction design, Human Factors, and requirements engineering, I hypothesized that the usability of interactive systems depends on their effectiveness, efficiency, and effort. I argued that in its turn, usability is one of the predictors of stakeholder satisfaction. I further stated that effectiveness relates to the degree of goal achievement with the system, that efficiency is based on the speed and accuracy of process execution and that effort has to do with the amount of energy invested in handling or understanding the system. Stakeholders supposedly require the maximum degree of goal achievement, as fast as possible and at the smallest possible error rate, against an optimal level of effort. Features of the machine and of the human side of an interactive system are assessed for their relevance (importance) and valence (expected gains or losses) towards these requirements. Complex evaluations can occur if the judgments about the machine do not run parallel to those about the human operator.

Several claims were corroborated whereas other were rejected by evidence from 928 employees of a multinational bank in The Netherlands, who interacted with 25 different banking systems (Study 1), as well as from 32 anesthetists working in the OR of a Dutch academic medical center (Study 2).

First, usability was not a central concern (cf. Jokela et al., 2003) in the minds of the investigated stakeholders and not always a prerequisite of stakeholder satisfaction (cf. Stakeholder Logistics). In Study 1, usability was not so much explained by satisfaction (cf. ISO 9241-11) but rather the other way round. In Study 2, usability and satisfaction were even unrelated. At best (Study 1), usability served as a mediator between performance factors on the one hand and satisfaction on the other as predicted by Stakeholder Logistics. At worst (Study 2), usability was merely a by-product of the efficiency of the

machines (refuting Stakeholder Logistics). In both cases, satisfaction and not usability was the end-product, which confirms Stakeholder Logistics.

Second, as an alternative, performance factors *were* important to predict usability as well as satisfaction measures. Confirming ISO 9241-11 and in this line, Stakeholder Logistics, effectiveness and efficiency contributed to usability in Study 1 and efficiency (of machines) in Study 2. In both cases, there were even direct effects of performance factors (i.e. efficiency) on satisfaction, thereby pushing usability on the side. As an additional corroboration of Stakeholder Logistics, the performance factors in Study 1 were relatively independent, being correlated but still retaining individual explanatory power.

Third, although performance factors and particularly efficiency were important to stakeholder satisfaction, unexpectedly, the judgments about human performance were completely separated from those about machines. This stresses the importance of making a distinction between humans and machines (cf. Seddon et al., 1998; Novick, 1997) when assessing performance and satisfaction estimates.

Fourth, in both cases, efficiency was the navel of stakeholder concerns, in particular, whether machines were error-free or not. As expected by the Efficiency model (Figure 3), efficiency was process oriented (Drucker 1954; 1974; Love, 1991). Efficiency judgments were not a mere reflection of time and speed aspects (ISO 9241-11; Frøkjær et al., 2000); in both cases, it consisted of time in coalition with accuracy estimates as predicted by the Efficiency model (cf. Ulrich & Hebert, 1982; Oviatt et al., 2004; Käki, 2004). Indeed, measures of particularly *inaccuracy* were most informative to assess efficiency judgments. It explained the estimated inefficiency of PCs and systems in Study 1 and judgments on machine as well as human efficiency in Study 2. In other words, errors are not part of effectiveness as ISO 9241-11 posits and capitalizing on speed in models of process optimization is a little beside the point because resolving inaccuracy is more important to stakeholders. As Nielsen and Levy (1994) found in their meta-analysis, increasing process speed does not necessarily lead to higher stakeholder satisfaction.

Fifth, as expected by the Efficiency model, it is worthwhile to decompose efficiency judgments not only into ‘efficient’ and ‘inefficient’ but also into judgments of fast vs. slow and accurate vs. inaccurate. In Study 1, the slow and inaccurate measures explained inefficiency of the banking system. In Study 2, fast and inaccurate explained human efficiency. In addition, estimates of machine speed and machine inaccuracy were capable of predicting the level of agreement to must requirements. In Study 2, the speed-accuracy trade-off usually observable in reaction-time paradigms (e.g., Usability Glossary, 2005; Oviatt et al., 2004; Meyer et al., 1988) was also retrievable in subjective judgments on human and machine efficiency. Here too, measures were partially correlated but yet maintained predictive power, thus indicating partial independence, as expected.

Sixth, the goals-to-requirements chiasm discovered in Chapter 4 could again be retrieved in the hospital data. This framework explains that requirements of a system are more vulnerable to change when goals and requirements hold opposite polarities in the stakeholder's mind (won't requirements that result from goals to achieve, must requirements that result from goals to avoid). In Study 2, the won't requirements of the OR monitoring system were predicted by human accuracy (goal to achieve). The must requirements were predicted by machine inaccuracy (a goal to avoid).

In sum, when it comes to which requirements of interactive systems change most, it is all about error reduction. Organizational managers should think up procedures to minimize human failure. On the side of the machines, debugging will enhance stakeholder satisfaction most. In system development, then, Software Engineers go before Usability Designers! This position counters Edberg and Olfman (2001), who report evidence that change requests on maintenance work were for 60% directed at the functionality and 15% of the change requests were directed at repair work such as debugging (cf. Chapter 1, Section 1.3.3). However, this 15% of error repair seems to impact the subjective experience of efficiency (and usability for that matter) more than the 60% of poor functionality. This links up nicely with the conclusions of Van Vugt et al. (2006). They report evidence that fluent game AI of the Sims2 game strongly affected involvement with the character, whereas the aesthetics of the graphic design did not affect the willingness to use it. In other words, good programming topples a nice interface. Efficiency requirements go before usability.

References

ANSI/INCITS 354 (2001). *Common Industry Format for Usability Test Reports.* InterNational Committee for Information Technology Standards (formerly NCITS).

Barnard, C. I. (1938). *The Functions of the Executive.* Cambridge, MA: Harvard University Press.

Baron, R., & Kenny, D. A. (1986). The moderator-mediator variabele distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51, 6,* 1173-1182.

Bevan, N. (2001). International standards for HCI and usability. *International Journal of Human-Computer Studies, 55,* 533-552.

Bevan, N., Kirakowski, J., & Maissel, J. (1991). What is usability? *Proceedings of the 4th International Conference on HCI, Stuttgart.* Amsterdam: Elsevier.

Brooke, J., Bevan, N., Brigham, F., Harker, S., & Youmans, D. (1990). Usability statements and standardisation - work in progress in ISO. *Human Computer Interaction – INTERACT '90.* Amsterdam: Elsevier.

Cooper, W. S. (1968). Expected search length: A single measure of retrieval effectiveness based on the weak ordering action of retrieval systems. *Journal of American Society of Information Science, 19, 1,* 30-41.

Dix, A. Finlay, J., Abowd, G., & Beale, R. (1998). *Human-Computer Interaction.* Hertfordshire, UK: Prentice Hall.

Drucker, P. F. (1954). *The Practice of Management.* New York: Harper and Row.
Drucker, P. F. (1974). *Management, Tasks, Responsibilities, Practices.* New York: Harper and Row.
Eason, K. (1988). *Information Technology and Organisational Change.* Bristol, PA: Taylor and Francis/Hemisphere Publishing.
Edberg, D., & Olfman, L. (2001). Organizational learning through the process of enhancing information systems. *34th Hawaii International Conference on System Sciences,*IEEE,http://csdl.computer.org/comp/proceedings/hicss/2001/0981/04/09814025.pdf.
Fitts, P. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychological, 47,* 381-391.
Frijda, N. H. (1986). *The Emotions.* New York: Cambridge University Press.
Frøkjær, E., Hertzum, M., & Hornbæk. K. (2000). Measuring usability: Are effectiveness, efficiency, and satisfaction really correlated? *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, The Hague, NL,* (pp. 345-352).
Hamilton, S., & Chervany, N. (1981). Evaluating information systems effectiveness – part 1: Comparing evaluation approaches. *MIS Quarterly, 5, 3,* 55-69.
Heitmiller, E. M. D. (2005). *How Anesthesia Works.* Howstuffworks.com. Retrieved August 3, 2005 from http://health.howstuffworks.com/anesthesia.htm.
Hoorn, J. F. (2005a). Sources of requirements change. A goal and viewpoints-driven explanation. *SIGCHI.NL* The Hague, Netherlands, October 13, 2005. New York: ACM Press.
Hoorn, J. F. (2005b). *ITO 01+02+03+04* (Tech. Rep.) [CD]. Computer Science, Information Management and Software Engineering. Amsterdam: Free University.
Hoorn, J. F., & Klaasse, B. (2006). *Efficiency Experiences in an Academic Medical Center* (Tech. Rep.) [CD]. Amsterdam: Free University.
ISO (1991). *International Standard ISO/IEC 9126. Information technology -- Software product evaluation - Quality characteristics and guidelines for their use,* International Organization for Standardization, International Electrotechnical Commission, Geneva.
ISO (1998). *Ergonomic Requirements for Office Work with Visual Display Terminals (VDTs) - Part 11: Guidance on Usability.* ISO 9241-11.
Jokela, T., Iivari, N., Matero, J., & Karukka, M. (2003). The standard of user-centered design and the standard definition of usability: analyzing ISO 13407 against ISO 9241-11. *Proceedings of the Latin American conference on Human-computer interaction Rio de Janeiro, Brazil* (pp. 53-60). New York: ACM Press.
Jordan, P. W. (1998). *An Introduction to Usability.* Bristol, PA: Taylor and Francis.
Käki, M. (2004). Proportional search interface usability measures. *NordiCHI '04, Tampere, Finland.*
Lauesen, S. (2005). *User Interface Design. A Software Engineering Perspective.* Harlow, UK: Pearson.
Love, A. J. (1991). *Internal Evaluation: Building Organizations from Within.* Newbury Park, CA: Sage.
MacCallum, R. C., Zhang, S., Preacher, K. J., & Rucker, D. D. (2002). On the practice of dichotomization of quantitative variables. *Psychological Methods, 7, 1,* 19-40.

Meyer, D. E., Irwin, D. E., Osman, A. M., & Kounios, J. (1988). The dynamics of cognition and action: Mental processes inferred from speed-accuracy decomposition. *Psychological Review, 95,* 183-237.

Newman, W. M., & Lamming, M. G. (1995). *Interactive System Design.* Reading, MA: Addison-Wesley.

Nielsen, J. (1993). *Usability Engineering.* San Diego, CA: Academic Press.

Nielsen, J., & Levy, J. (1994). Measuring usability: Preference vs. performance. *Communications of the ACM, 37, 4,* 66-75.

Novick, D. G. (1997). What is Effectiveness? *Position paper for the CHI '97 Workshop on HCI Research and Practice Agenda Based on Human Needs and Social Responsibility Atlanta GA.*

Oviatt, S., Coulston, R., & Lunsford R. (2004). When do we interact multimodally?: Cognitive load and multimodal communication patterns. *Proceedings of the 6th International Conference on Multimodal Interfaces* (pp. 129-136). New York: ACM Press.

Pather, S., Erwin, G., & Remenyi, D. (2003). Measuring e-commerce effectiveness: A conceptual model. *Proceedings of the 2003 Annual Research Conference of the South African Institute of Computer Scientists and Information Technologists on Enablement through Technology* (pp. 143-152). South African Institute for Computer Scientists and Information Technologists. Republic of South Africa.

Preacher, K. J., & Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior Research Methods, Instruments, & Computers, 36, 4,* 717-731.

Preece, J., Rogers, Y., & Sharp, H. (2002). *Interaction Design: Beyond Human-Computer Interaction.* New York: Wiley.

Scapin, D. L., & Bastien, J. M. C. (1997). Ergonomic criteria for evaluating the ergonomic quality of interactive systems. *Behaviour & Information Technology, 16, 4/5,* 220-231.

Scott, J. E. (1995). The measurement of information systems effectiveness: Evaluating a measuring instrument. *ACM SIGMIS Database, 26, 1,* 43-61.

Seddon, P. B., Staples, D. S., Patnayakuni, R., & Bowtell, M. J. (1998). The IS effectiveness matrix: The importance of stakeholder and system in measuring IS success. *Proceedings of the International Conference on Information Systems, Helsinki, Finland* (pp. 165-176).

Shackel, B. (1991). Usability – context, framework, definition, design and evaluation. *Human Factors for Informatics Usability* (pp. 21-37). New York: Cambridge University Press.

Shneiderman, B. (1998). *Designing the User Interface.* Reading, MA: Addison-Wesley.

Silverstein, C., Henzinger, M., Marais, J., & Moricz, M. (1998). *Analysis of a Very Large Alta Vista Query Log* (Tech. Rep. 1998-014). Palo Alto, CA: COMPAQ Systems Research Center.

Sobel, M. E. (1982). Asymptotic intervals for indirect effects in structural equations models. In S. Leinhart (Ed.), *Sociological Methodology 1982* (pp. 290-312). San Francisco, CA: Jossey-Bass.

Ulrich, E., & Hebert, D. (1982). Speed and accuracy in digital network simulation based on structural modeling. *Proceedings of the 19th Conference on Design Automation* (pp. 587-593). Piscataway, NJ: IEEE Press.

Usability Glossary (2005). *Speed-accuracy Tradeoff*. Retrieved March 09, 2005 from http://www.usabilityfirst.com/ glossary/term_475.txl

Van Vugt, H. C., Hoorn, J. F., Konijn, E. A., & De Bie Dimitriadou, A. (2006). Affective affordances: Improving interface character engagement through interaction. *International Journal of Human-Computer Studies, 64, 9,* 874-888.

Van Welie, M., Van der Veer, G., & Eliëns, A. (1999). Breaking down usability. *Proceedings of Interact '99* (pp. 613-620), August 30 - September 3, Edinburgh, Scotland.

Wambugu, C. K. (1982). The library manager. *Reader of the Unesco International Symposium on the Harmonization of Education and Training Programmes in Information Science, Karfa, Nairobi* (Chapter 10).

Wickens, C. D., Lee, J. D., Liu, Y., & Gordon Becker, S. E. (2004). *An Introduction to Human Factors Engineering.* Upper Saddle River, NJ: Pearson.

Yerkes R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology and Psychology, 18,* 459-482.

Appendix 5.1

OR Efficiency Questionnaire[9]

De volgende vragen gaan over de computer waarop de markeerpunten worden ingevoerd. Voorbeelden van markeerpunten zijn 'begin inleiding' en 'begin incisie'.

Ik vind:

		Helemaal mee oneens	Mee oneens	Beetje mee oneens	Beetje mee eens	Eens	Helemaal mee eens
1.	Dat de markeerpunten op meer computers bijgehouden moeten worden.	0	1	2	3	4	5
2.	Dat de computer een bevestiging moet vragen bij een afwijkende invoer.	0	1	2	3	4	5
3.	Dat je via de computer moet kunnen aangeven wanneer de premedicatie gegeven moet worden.	0	1	2	3	4	5
4.	Dat je meer markeerkeerpunten moet bijhouden dan nu.	0	1	2	3	4	5

Ik vind:

		Helemaal mee oneens	Mee oneens	Beetje mee oneens	Beetje mee eens	Eens	Helemaal mee eens
5.	Dat je moet kunnen aangeven dat een zitting uitloopt.	0	1	2	3	4	5
6.	Dat de computer automatisch een schatting van de zittingsduur moet geven.	0	1	2	3	4	5
7.	Dat de computer om de vijf minuten een invoerwaarschuwingspiep moet geven.	0	1	2	3	4	5
8.	Dat je een patiënt moet kunnen bestellen met de computer.	0	1	2	3	4	5

Pagina 2 van 8

[9] The first page is missing because it contains strategic information about the organization.

Geef uw oordeel over de handigheid in gebruik door u of door anderen van de volgende apparaten.

De volgende apparaten zijn handig in het gebruik:

		Helemaal mee oneens	Mee oneens	Beetje mee oneens	Beetje mee eens	Eens	Helemaal mee eens
9.	De monitor van vitale functies	0----------1---------2---------3---------4---------5					
10.	De verdamper	0----------1---------2---------3---------4---------5					
11.	De infuuspomp	0----------1---------2---------3---------4---------5					
12.	Het beademingsapparaat	0----------1---------2---------3---------4---------5					
		Helemaal mee oneens	Mee oneens	Beetje mee oneens	Beetje mee eens	Eens	Helemaal mee eens
13.	Het röntgenapparaat	0----------1---------2---------3---------4---------5					
14.	De AIS invoer module	0----------1---------2---------3---------4---------5					
15.	De coagulator	0----------1---------2---------3---------4---------5					
16.	Het afzuigapparaat	0----------1---------2---------3---------4---------5					

(Deel 1)
Er volgen drie delen met vragen over het werken met machines in de operatiekamer. Dit is het eerste deel.

Geef aan in welke mate u het eens bent met de volgende stellingen:

		Helemaal mee oneens	Mee oneens	Beetje mee oneens	Beetje mee eens	Eens	Helemaal mee eens
17.	De bewakingsapparatuur is binnen een mum van tijd aangesloten.	0	1	2	3	4	5
18.	Het dichten van de operatiewond is een traag proces.	0	1	2	3	4	5
19.	De zelftest van het anesthesieapparaat werkt correct.	0	1	2	3	4	5
20.	Bij het aansluiten van de beademingsapparatuur treden fouten op.	0	1	2	3	4	5
		Helemaal mee oneens	Mee oneens	Beetje mee oneens	Beetje mee eens	Eens	Helemaal mee eens
21.	Het positioneren van een patiënt gebeurt zorgvuldig.	0	1	2	3	4	5
22.	De AIS invoer module heeft een lange opstarttijd.	0	1	2	3	4	5
23.	De beademingsapparatuur werkt ogenblikkelijk.	0	1	2	3	4	5
24.	Het lukt slecht gegevens in te voeren in het AIS.	0	1	2	3	4	5

Pagina 4 van 8

De volgende stellingen vragen een algemeen oordeel over de zitting, personeel en apparatuur. Geef aan in hoeverre de volgende zaken <u>inefficiënt</u> verlopen.

Ik vind de volgende zaken inefficiënt:

		Helemaal mee oneens	Mee oneens	Beetje mee oneens	Beetje mee eens	Eens	Helemaal mee eens
25.	Het verloop van een zitting.	0----------1---------2---------3---------4---------5					
26.	De werkwijze van het OK personeel.	0----------1---------2---------3---------4---------5					
27.	De ondersteuning door de apparaten.	0----------1---------2---------3---------4---------5					

Geef ook aan hoe <u>tevreden</u> u bent over de volgende zaken.

Ik ben tevreden over:

		Helemaal mee oneens	Mee oneens	Beetje mee oneens	Beetje mee eens	Eens	Helemaal mee eens
28.	Het verloop van een zitting.	0----------1---------2---------3---------4---------5					
29.	De werkwijze van het OK personeel.	0----------1---------2---------3---------4---------5					
30.	De ondersteuning door de apparaten.	0----------1---------2---------3---------4---------5					

Pagina 5 van 8

(Deel 2)
Nu volgt het tweede deel met stellingen over het werken van mensen met machines in de operatiekamer.

Geef aan in welke mate u het eens bent met de volgende stellingen:

		Helemaal mee oneens	Mee oneens	Beetje mee oneens	Beetje mee eens	Eens	Helemaal mee eens
31.	De röntgenanalist is lang bezig met het maken van een foto.	0	1	2	3	4	5
32.	Het dichtbranden met de coagulator is gauw gebeurd.	0	1	2	3	4	5
33.	De instrumenten zijn op de juiste manier klaargelegd.	0	1	2	3	4	5
34.	De zelftest van het anesthesieapparaat is tijdrovend.	0	1	2	3	4	5
		Helemaal mee oneens	Mee oneens	Beetje mee oneens	Beetje mee eens	Eens	Helemaal mee eens
35.	Het inzetten van een spuit in de infuuspomp gaat mis.	0	1	2	3	4	5
36.	De beademingsapparatuur werkt foutenvrij.	0	1	2	3	4	5
37.	De bewakingsmonitor geeft informatie verkeerd door.	0	1	2	3	4	5
38.	Het controleren van de apparatuur door de anesthesist duurt kort.	0	1	2	3	4	5
		Helemaal mee oneens	Mee oneens	Beetje mee oneens	Beetje mee eens	Eens	Helemaal mee eens
39.	De coagulator werkt precies.	0	1	2	3	4	5
40.	De AIS invoer module is traag tijdens het invoeren van gegevens.	0	1	2	3	4	5
41.	Het klaarzetten van de röntgenapparatuur heeft tijd nodig.	0	1	2	3	4	5
42.	Het röntgenapparaat is foutgevoelig.	0	1	2	3	4	5
		Helemaal mee oneens	Mee oneens	Beetje mee oneens	Beetje mee eens	Eens	Helemaal mee eens
43.	Het automatisch registreren van de anesthesiegegevens mislukt.	0	1	2	3	4	5
44.	Het chirurgisch team gaat secuur om met hun steriliteit.	0	1	2	3	4	5
45.	De röntgenanalist maakt vergissingen tijdens het nemen van foto's.	0	1	2	3	4	5
46.	De verdamper is meteen te gebruiken.	0	1	2	3	4	5

Pagina 6 van 8

De volgende stellingen vragen weer een algemeen oordeel over de zitting, personeel en apparatuur.
Geef aan in hoeverre de eerder genoemde zaken <u>efficiënt</u> zijn.
(U hoeft niet consequent te zijn.)

Ik vind de volgende zaken efficiënt:

		Helemaal mee oneens	Mee oneens	Beetje mee oneens	Beetje mee eens	Eens	Helemaal mee eens
47.	Het verloop van een zitting.	0	1	2	3	4	5
48.	De werkwijze van het OK personeel.	0	1	2	3	4	5
49.	De ondersteuning door de apparaten.	0	1	2	3	4	5

Geef ook aan hoe <u>ontevreden</u> u bent over de eerder genoemde zaken.
(U hoeft niet consequent te zijn.)

Ik ben ontevreden over:

		Helemaal mee oneens	Mee oneens	Beetje mee oneens	Beetje mee eens	Eens	Helemaal mee eens
50.	Het verloop van een zitting.	0	1	2	3	4	5
51.	De werkwijze van het OK personeel.	0	1	2	3	4	5
52.	De ondersteuning door de apparaten.	0	1	2	3	4	5

Pagina 7 van 8

(Deel 3)
Dit is het laatste deel en gaat weer over het werken van mensen met machines in de operatiekamer.

Geef aan in welke mate u het eens bent met de volgende stellingen:

		Helemaal mee oneens	Mee oneens	Beetje mee oneens	Beetje mee eens	Eens	Helemaal mee eens
53.	Mensen laten spullen uit hun handen vallen.	0	1	2	3	4	5
54.	Het dichten van de operatiewond gebeurt nauwkeurig.	0	1	2	3	4	5
55.	Het positioneren van een patiënt is zo gebeurd.	0	1	2	3	4	5
56.	Het handen wassen door het chirurgisch team is een langdurige procedure.	0	1	2	3	4	5
		Helemaal mee oneens	Mee oneens	Beetje mee oneens	Beetje mee eens	Eens	Helemaal mee eens
57.	Het röntgenapparaat laat lang op zich wachten.	0	1	2	3	4	5
58.	De bewakingsmonitor is direct bruikbaar.	0	1	2	3	4	5
59.	De verdamper werkt op juiste wijze.	0	1	2	3	4	5
60.	De patiënt is vlug afgedekt.	0	1	2	3	4	5

Appendix 5.2

OR Equipment

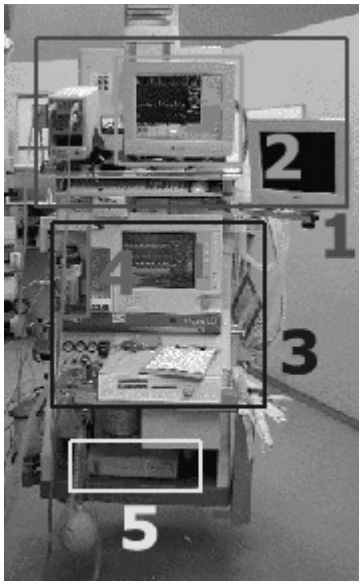
Anesthesia device

This device is placed at the head-end of the operation table and consists of a number of machines placed in a rack on wheels that is connected to the ceiling with flex. Five devices are placed on this rack (1-5).

Marquette input module (1)

Registers general OR information, which can be entered automatically or manually by the anesthetist.

Vital signs monitor (2) Displays the values of the vital functions of the patient: ECG, pulse rate, temperature, non-invasive and invasive blood pressure, and respiration. Values are indicated by color, waveform, and/or number. The monitor warns when values exceed a certain criterion.



Anesthesia device (photo Tijink, 2005).

Respiratory ventilator of Dräger (3) When in anesthesia, the respiratory ventilator provides the patient artificial respiration. The display shows CO_2 and oxygen value, the frequency and the amount of air per breath.

Vaporizer (4) Inserts a continuous flow of anesthesia. It can give accurate doses under different environmental circumstances.

Local Recorder (5) **R**ecords all information concerning the anesthesia device and the patient. OR personnel cannot interact with it and thus, this device is not part of my study.

The infusion pump

Several liquids are brought directly into the patient's veins. Some liquids are needed to keep up the body's homeostasis; other are given as an anesthesia. Because the latter has to be given in precise doses, it is inserted by an infusion pump. The syringe with drugs is attached to the infusion pump. The amount of medication per hour is set to give the right dosage.

Infusion pump

(http://www.wpiinc.com/World Precision Instruments)

Coagulator

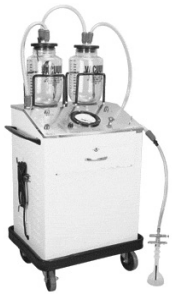
(http://www.barryhsu.com/)

Coagulator

Burning device to close the smaller and less important blood vessels. Large important vessels are constricted and later on reconnected. The coagulator is usually controlled with a foot pedal.

Suction device

Removes blood and cleans wounds before they are closed. The device is connected by tubes to a container where the liquids are collected. When full, the containers have to be replaced by hand.



Suction device

(http://www.suctioninfo.com)

X-ray machine

(http://www.arabtradezone.com/prophoto/53047/1079335672.gif)

X-ray machine

The X-ray device is not standard OR equipment and can be called if necessary. The device is large and has to be placed around the patient. All personnel in the OR wear a lead skirt to keep them from radiation damage.

Two screens show the photos. The X-ray specialist should make a sharp picture of what the surgeon wants to see. This takes a lot of adjusting. The X-ray device has a digital photo store.

6 Conclusions

6.1 Back to CoStaR

The starting point of the empirical research offered in the previous chapters was the Change of Stakeholder Requirements (CoStaR) model described in Chapter 2 (Section 2.10, Figure 7). What catches the eye is that many of the factors that were supposed to explain requirements change indeed were effective except for one, egotistic (selfish) vs. altruistic goals (Chapter 3 and 4). It was replaced by the factor of Model Change (Chapter 3), pointing out that the changes from old to new (or v.v.) personal or business models contribute to variability in the level of agreement with requirements. The other thing worth mentioning is that the *structure* of how these factors cooperate and affect one another was almost completely reorganized. This prompted the need to define several sub models that accounted for particular aspects of the CoStaR model. Figure 1 shows which factors of CoStaR survived and what their test status is.

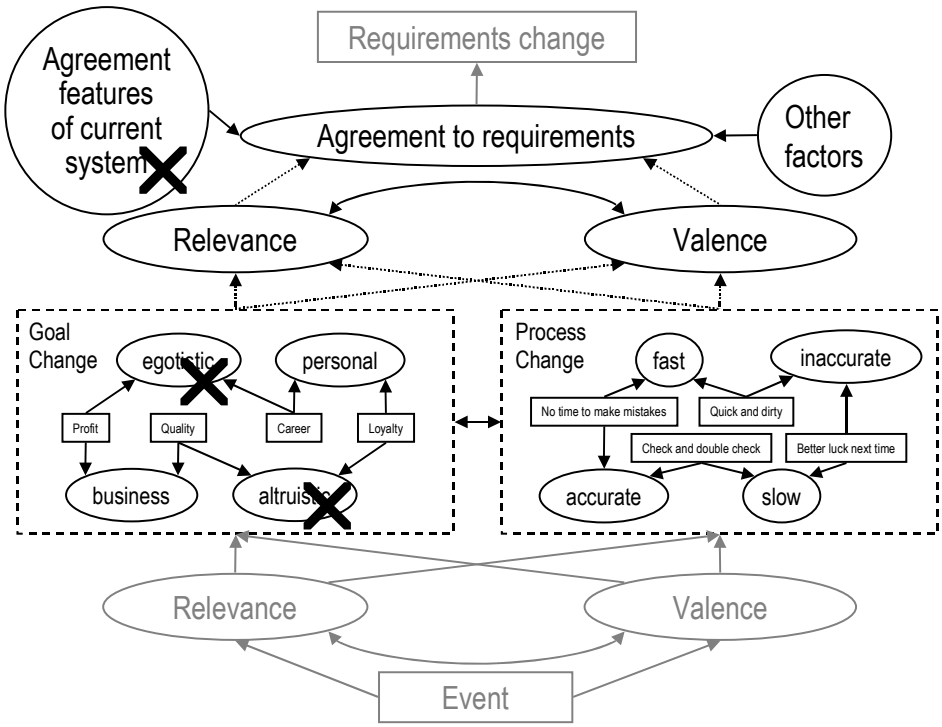


Figure 1: Change of Stakeholder Requirements model with test status. Crossed out – variables have no significant effect. Dotted lines – relations differ slightly from hypotheses. Gray lines – not tested.

The effects of events, at the bottom of Figure 1 (gray), were not tested. Chapter 2 (Section 2.11) explained that companies only allowed me to investigate

business model change *after* the event (e.g., a financial buy-out) had happened. Not before or during the event, because that concerned too critical business information. In other words, CoStaR can only hypothesize *that* events affect model change but not with any more precision and without substantial evidence.

The other point was that the actual change or requirements, at the top of Figure 1 (gray), was not investigated for the systems under development. This had to do with avoiding possible contamination of observations and with the different time lines of the business's system development and conducting scientific research (Chapter 2). The work in this book mainly concerned measuring (dis)agreement to requirements and advising the company on alternatives for their requirements list. If the company decided to actually change the requirements it was not distinguishable any more where the effect came from. It could be the effect of the opinion of those 'official' stakeholders who were not involved in actually designing the system (e.g., end-users). This was the thing that was actually measured. It could also be the effect of those 'official' stakeholders who were involved in designing the system as well as in conducting the research (e.g., in the case of the CMS and LWMS – Chapter 3 and 4). The roles of these designers/researchers were contaminated. Or requirements change could be the effect of the researchers' doings as unofficial stakeholders of the system who operated as external advisors (cf. Chisăliță et al., 2002). Wanting to control for this problem of the 'immersive' ethnographical approach to system design took too much time of the companies involved and so this line of investigation was canceled.

However, everything in between, the processes that supposedly lead from events to actually changing the requirements, was investigated thoroughly. This part was not contaminated by changing roles between designer-advisor and scientific investigator because it merely observed the state of mind of those 'official' stakeholders that were not designers without interfering (yet) as external advisor.

6.1.1 Variability in Agreement

The layer of Figure 1 that is dedicated to the "Agreement to requirements" is based on the insights of Kotonya and Sommerville (1989) that the requirements engineering process ultimately is directed at creating a list of requirements that are agreed-upon (Chapter 2, Figure 5). It is not too bold to assume, then, that changes proposed to this list originate from changes in agreement to the requirements concerned. By definition, then, requirements that are subjected to change are requirements that, in a sample of stakeholders, show more variability in the level of agreement than requirements that are stable. Put simply, they raise more discussion.

6.1.2 The Goals-to-Requirements Chiasm

This approach to measuring agreement with requirements opened the amazing vista of the goals-to-requirements chiasm as a possible explanation of requirements change (Chapter 4 and 5). First, "Agreement to requirements" is too general an approach. Requirements need to be decomposed at least into must versus won't requirements because the explanation why stakeholders (dis)agree with them differs. Second, requirements of a system are more sensitive to fluctuations in agreement (and following from that, to requirements change) when stakeholders connect requirements to goals of opposite polarity. That is, variability in agreement is high (agree vs. disagree) when won't requirements, which are negative, should cover goals to achieve, which are positive. An example could be that a Windows '95 operating system, a requirement to which a stakeholder may disagree, should guarantee system stability, a goal to which s/he agrees. Variability in agreement is also high when must requirements, which are positive, should serve goals to avoid, which are negative. For instance, automated debugging, a requirement to which a stakeholder may agree, should circumvent machine error, a machine state to which the stakeholder disagrees. Requirements that uphold 'straight' relationships (goals to approach with must requirements and goals to avoid with won't requirements) are the stable requirements. The goals and requirements in such 'straight' combinations raise either agreement or disagreement. For example, Unix (agree) for stability (agree). Windows '95 (disagree) for instability (disagree).

6.1.3 The Type of Goals

Where requirements change did not come from was (dis)agreement with the current system (Chapter 4, LWMS case), satisfaction with humans or machines at work (Chapter 5), usability of machines (Chapter 5), or the changes from egotistic (selfish) to altruistic goals (or v.v.) (Chapter 3 and 4). This means that "Agreement with features of the current system" and the variables egotistic and altruistic in Figure 1 can be omitted (black crosses).

What did matter was that personal goals overshadowed the effects of business goals and that efficiency was a central stakeholder concern. The box of Goal Change in Figure 1 should be revised. Although personal vs. business goals made a difference for agreement with requirements, for the relevance of goals, and for requirements prioritization (Chapter 3), egotistic vs. altruistic did not. Instead, changes from old to new goals (and v.v.) had a greater impact. Stakeholders were inclined to make more changes when they looked at a list of requirements from a personal viewpoint compared to a business viewpoint. Moreover, this effect increased when stakeholders acquired a new set of goals (i.e. personal goals) as compared to an older position such as an outdated business model.

Within the personal or business models, the type of goals that seemed to matter most was related to processes – i.e. boosting efficiency by becoming

less error-prone (Chapter 5). Requirements change, then, can be expected from performance requirements that are derived from lower-level goals (cf. Anton, Cracken, & Potts, 1994; Alves & Finkelstein, 2002) such as improving personal efficiency or optimizing a business process. In reflecting the goals-to-requirements chiasm, stakeholders wish to achieve personal goals (positive) with less inaccurate machines (thus making the negative less negative).

The double-headed arrow between the box of Process Change and the Goal Change box (Figure 1) is correct in that a desired state of a process (e.g., efficiency) can become a personal or business goal to achieve in the next version of a software product, thus predicating requirements change. In addition, the decomposition of processes in speed and accuracy components (Figure 1, Process Change box) was most fruitful, because these components (i.e. estimated inaccuracy) predicted the judgments on human and machine efficiency, which in their turn best predicted stakeholder satisfaction (Chapter 5). Moreover, human accuracy affected changes in agreement with won't requirements and machine inaccuracy with must requirements (Chapter 5).

6.1.4 Relevance-Valence

Relevance, the amount of importance stakeholders attach to a requirement in view of their goals and concerns, was crucial to understand changes in requirements prioritization (Chapter 3). It showed, for example (Chapter 3), that even when both business and personal goals are considered irrelevant, the personal goals are still considered more important (less irrelevant) than the business goals. In other words, a business system that does not take the personal needs of its stakeholders seriously is bound to fail (cf. Price & Cybulski, 2004).

The effects of Valence, the expected positive or negative outcome of using the requirements once implemented, was repeatedly established in Chapter 4. Valence added to the explanation of changes in agreement to requirements, not always as a vital step to arrive at agreement (valence as mediator), but often as an important influence 'from outside' (valence as moderator).

This changes the constellation depicted in Figure 1 (dotted lines). There, Valence is modeled as a mediator, which sometimes was true (i.e. for must requirements) but more often not (i.e. for won't requirements). Because the way Relevance was measured did not allow regression analysis, the status of Relevance as moderator is based on similarity with Valence rather than on rigorous tests.

6.1.5 The Requirements-analysis Rift

Chapter 3 and 4 pointed at the existence of the requirements-analysis rift. This is a viewpoints-dependent phenomenon (cf. Sommerville & Sawyer, 1997) that occurs when goals and requirements appear to be disconnected in the mind of the stakeholder. Stakeholders consider requirements as something of the business whereas the goals that these requirements preferably should

satisfy are personal (see previous section). If the connections between requirements and goals are not spelled out, stakeholders do not necessarily connect the management-viewpoint requirements to their personal goals. Requirements engineers should help switch the focus from a business to a personal point of view when requirements are negotiated. Because the rift occurs when requirements are perceived as unconnected to goals, by definition, the goals-to-requirements chiasm can only come into operation when the requirements-analysis rift is resolved.

6.2 Revised CoStaR

Figure 2 offers a theoretical integration of the various (sub) models derived from the empirical data. The requirements-analysis rift is represented by a switch symbol for normally open contact (–/ –). The mechanism that explains requirements change, the goals-to-requirements chiasm (mid part Figure 2), is out of order if the requirements engineer does not connect business requirements to personal goals. Once connected, two conditions play a major role in the way the chiasm progresses.

Figure 2 shows two changer or converter symbols, in which the input and output are inserted in each half of the general symbol (box with diagonal) to show the nature of change. This may be from the old business model (BM1) to the new (BM2) or v.v. Or, the change may be from an old (PM1) to a new personal model (PM2) or v.v. In Figure 2, Model Change (converter symbols) determine the Goals to Approach (e.g., maintain human accuracy) as well as Goals to Avoid (e.g., machine inaccuracy).

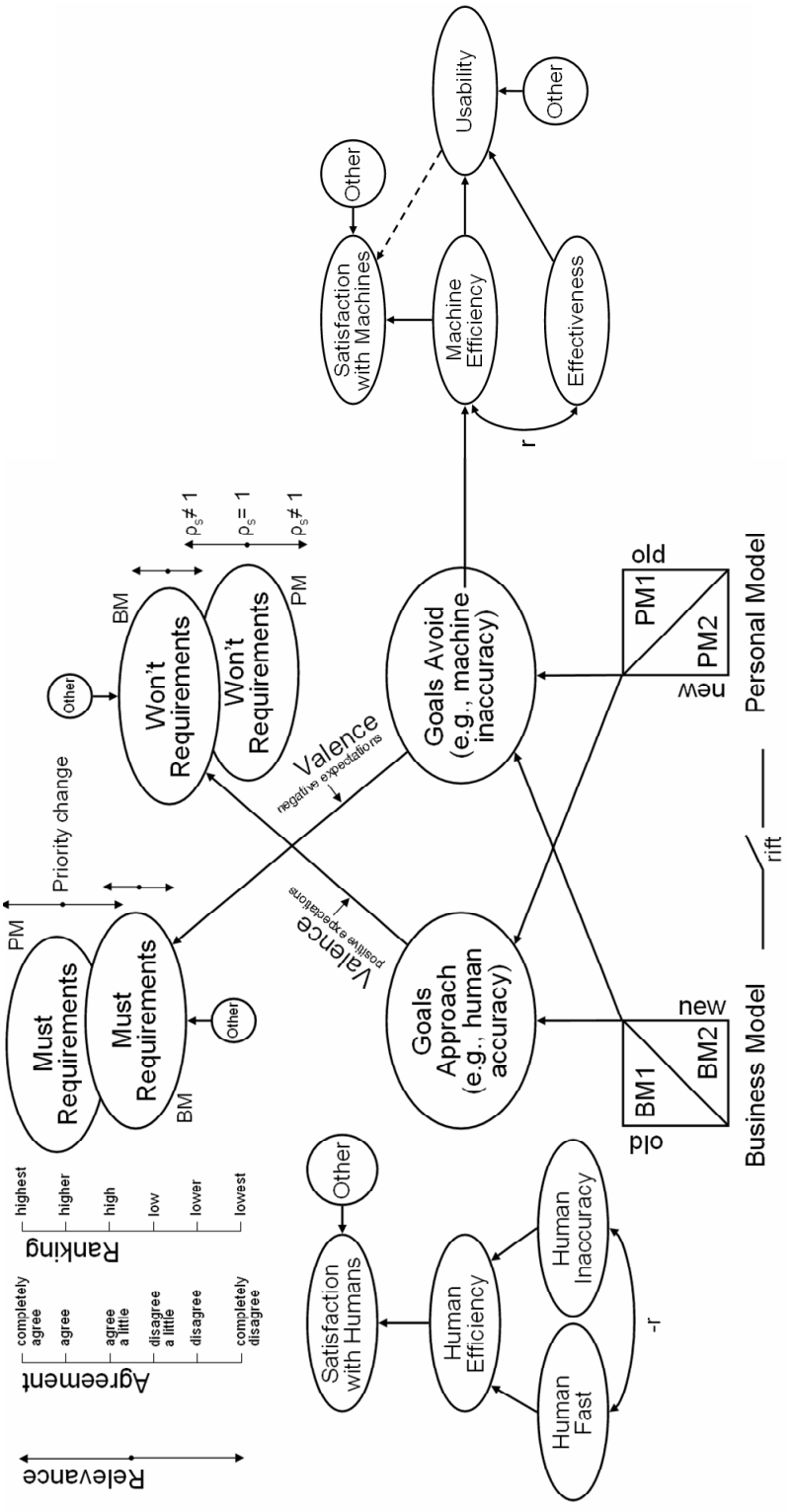


Figure 2: Revised CoStaR model with extensions to Stakeholder Logistics. 'Other' refers to factors that may account for unexplained variance.

Goals to Avoid determine the changeability in agreement with Must Requirements. Goals to Approach determine the changeability in disagreement with Won't Requirements (the chiasm). Valence, in how far stakeholders expect that a requirement is helpful or obstructive, moderates the level of agreement with requirements. Positive expectations affect the relation between Approach and Won't; negative expectations between Avoid and Must. In the first case, this means that stakeholders expect to accomplish a goal brilliantly if certain features are taken off the system (e.g., an RSI warning add-on that interrupts a person's work all the time). In the second case, if a LaserJet from the 1010/1012 series is a must because of its speed; stakeholders may expect trouble with regard to avoiding machine error because it is known to have a firmware bug. This last example also illustrates why Machine Speed (contributing to Must Requirements – Chapter 5) is left out of Figure 2; that variable probably added more to requirements stability (all systems must always be fast) rather than change (the danger of inaccuracy makes the requirement controversial).

The upper part of the mid section in Figure 2 actually shows two extra ellipses for Must and Won't Requirements. This is to illustrate that the chiasm happens at a higher level of relevance for the stakeholder's personal goals than for the business goals. A must that serves a personal goal is far more relevant to a stakeholder than that same requirement serving a business goal. Likewise, a won't requirement that damages a personal goal is disagreed to more than when it damages a business objective.

The Relevance scale at the far left of Figure 2 has a black dot that represents scale base-point. From there, Relevance can boost either agreement or disagreement with a requirement. Subsequently, this may change the rank order of the requirement in the priority list. Because the level of Relevance underlies the prioritization of requirements (Chapter 3), the size or amplitude of the change in priorities is bigger under conditions of personal change than under conditions of business-model change. Therefore, the straight double-headed arrows next to the Must and Won't ellipses are larger for the personal condition (PM) than for the business condition (BM). The larger the difference (e.g., from highest to lowest ranking), the more Spearman's rho (ρ_s) as an indicator of priority change moves away from 1 (no change). This accounts for the repeated finding in Chapter 3, that stakeholders showed more priority change (ρ_s approaching 0) when personal models changed than when business models did.

On the right hand-side of Figure 2, you find the variables in the speed-accuracy trade-off that are relevant for agreement with requirements, stakeholder satisfaction, and usability. As a pre-eminent example, machine inaccuracy is a Goal to Avoid. Figure 2 shows that estimated inaccuracy of machines not only may change agreement to the Must Requirements, it directly affects judgments about Machine Efficiency as well, which is the most important contributor to the stakeholders' Satisfaction with the Machines. Together with

estimated Effectiveness (Chapter 5, Study 1), moreover, do judgments about Machine Efficiency predict the level of Usability. Usability sometimes contributes to Satisfaction (Chapter 5, Study 1) and at other times it does not (Chapter 5, Study 2) (Figure 2, dashed arrow).

Completely separated from the machine side of a to-be-developed interactive system are the judgments about Human Efficiency (Figure 2, sub model on the far left). Although human accuracy is a good example of a Goal to Approach, Satisfaction with how Humans perform their jobs is seen as fully independent of the machines they work with. From a CHI point of view, here the stakeholders make a bit of an overstatement. Human Efficiency is the main predictor of the level of Satisfaction with Humans and this efficiency is based on the trade-off between people working Fast while trying to avoid Inaccuracy.

In most of the studies I performed, stakeholders were preoccupied with improving the efficiency of the current system. The main conclusion of this thesis, then, is that we need *user-centered design for machine-centered users*.

6.3 Back to DUTCH

In Chapter 2, Section 2.4, the design approach called DUTCH (Designing for Users and Tasks from Concepts to Handles) was supplemented with business models. This was to enrich the task model with business requirements so to enhance the similarity with a full-fledged requirements specification. However, the research presented in this book urges to include personal models as well (Figure 3).

The task hierarchy of Business Model 1 (BM1) corresponds to the task or process hierarchy in TM1. Yet, both should be in alignment with the personal goals and work processes in Personal Model 1 (PM1). The task hierarchy of Business Model 2 (BM2) corresponds to the task or process hierarchy in TM2. To anticipate requirements change in an early stage, however, again both should be in accordance with the newly acquired personal goals and work processes in Personal Model 2 (PM2).

BM1 relates business requirements to TM1 (goals, Critical Success Factors, problems) but TM1 should also take personal requirements (PM1) into account. BM1, PM1, and TM1 are analyzed, leading to TM2. The task hierarchy of TM2 is used to build BM2 but also PM2, which relates the business and personal requirements to the desired situation. BM2 illustrates how business goals, CSFs, and problems affect TM2 and serves as a justification of TM2 to that extent. The same should be done for PM2. If this justification fails, the original DUTCH model did not offer the possibility to change the requirements. By introducing the CoSTaR module (Figure 3), it is possible to redesign the system in a timely manner.

Once there is a (re)specification of the User Virtual Machine with NUAN models, the implementation can start. And so we have arrived at requirements management that on the one hand provides a specification of the requirements

that are agreed-upon by the stakeholders and on the other can account for requirements change.

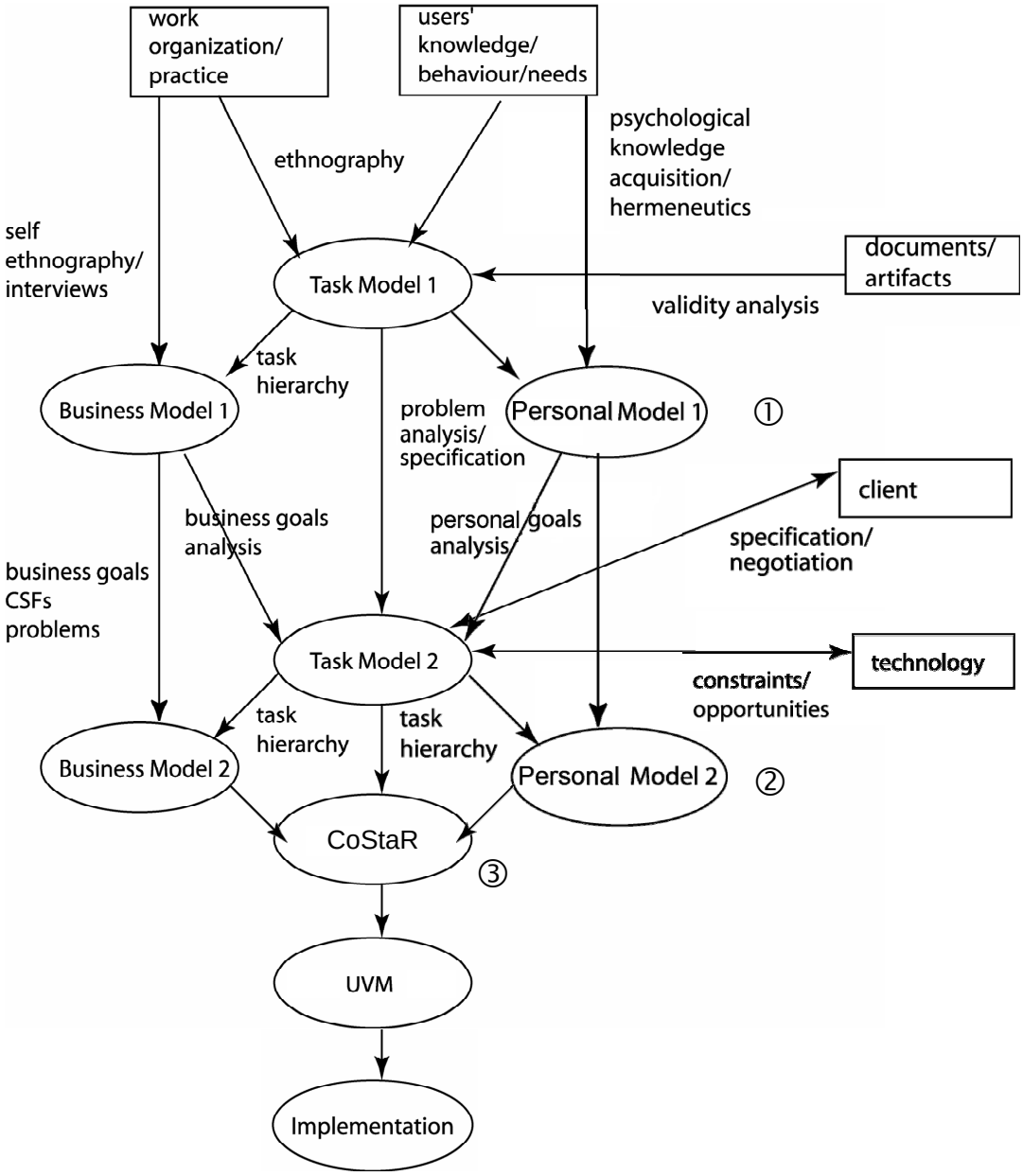


Figure 3: The DUTCH approach with business and personal model extensions (① and ②) as well as a module to account for requirements change (CoStaR) ③.

References

Alves, C., & Finkelstein, A. (2002). Challenges in COTS-decision making: A goal-driven requirements engineering perspective. *Workshop on Software Engineering Decision Support* (Ischia IT, 2002).

Anton, A.I., Cracken, W., & Potts, C. (1994). Goal decomposition and scenario analysis in business process reengineering, in *Proceedings of CAiSE '94* (Utrecht NL, 1994).

Chisăliță, C. M., Van der Veer, G. C., Hoorn, J. F., & Puerta Melguizo, M. C. (2002). A model of cultural influences on team work practice - a design team case-study. In S. Bagnara, S. Pozzi, A. Rizzo, & P. Wright (Eds.), *Proceedings of the Eleventh European Conference on Cognitive Ergonomics - Cognition, Culture, and Design* (pp. 151-156). Catania, Italy: ECCE.

Kotonya, G., & Sommerville, I. (1998). *Requirements Engineering: Processes and Techniques.* New York, NY: John Wiley.

Price, J., & Cybulski, J. (2004). Influence of stakeholder communication on consensus making in requirements negotiation. In *Proceedings of AWRE*, Adelaide, Australia.

Sommerville, I., & Sawyer, P. (1997). Viewpoints: Principles, problems and a practical approach to requirements engineering. *Annals of Software Engineering 3*: 101-130.

SIKS Dissertation Series

====
1998
====

1998-1 Johan van den Akker (CWI)
DEGAS - An Active, Temporal Database of Autonomous Objects

1998-2 Floris Wiesman (UM)
Information Retrieval by Graphically Browsing Meta-Information

1998-3 Ans Steuten (TUD)
A Contribution to the Linguistic Analysis of Business Conversations within the Language/Action Perspective

1998-4 Dennis Breuker (UM)
Memory versus Search in Games

1998-5 E.W.Oskamp (RUL)
Computerondersteuning bij Straftoemeting

====
1999
====

1999-1 Mark Sloof (VU)
Physiology of Quality Change Modelling; Automated modelling of Quality Change of Agricultural Products

1999-2 Rob Potharst (EUR)
Classification using decision trees and neural nets

1999-3 Don Beal (UM)
The Nature of Minimax Search

1999-4 Jacques Penders (UM)
The practical Art of Moving Physical Objects

1999-5 Aldo de Moor (KUB)
Empowering Communities: A Method for the Legitimate User-Driven Specification of Network Information Systems

1999-6 Niek J.E. Wijngaards (VU)
Re-design of compositional systems

1999-7 David Spelt (UT)
Verification support for object database design

1999-8 Jacques H.J. Lenting (UM)
Informed Gambling: Conception and Analysis of a Multi-Agent Mechanism for Discrete Reallocation.

====
2000
====

2000-1 Frank Niessink (VU)
Perspectives on Improving Software Maintenance

2000-2 Koen Holtman (TUE)
Prototyping of CMS Storage Management

2000-3 Carolien M.T. Metselaar (UVA)
Sociaal-organisatorische gevolgen van kennistechnologie; een procesbenadering en actorperspectief.

2000-4 Geert de Haan (VU)
ETAG, A Formal Model of Competence Knowledge for User Interface Design

2000-5 Ruud van der Pol (UM)
Knowledge-based Query Formulation in Information Retrieval.

2000-6 Rogier van Eijk (UU)
Programming Languages for Agent Communication

2000-7 Niels Peek (UU)
Decision-theoretic Planning of Clinical Patient Management

2000-8 Veerle Coup, (EUR)
Sensitivity Analyis of Decision-Theoretic Networks

2000-9 Florian Waas (CWI)
Principles of Probabilistic Query Optimization

2000-10 Niels Nes (CWI)
Image Database Management System Design Considerations, Algorithms and Architecture

2000-11 Jonas Karlsson (CWI)
Scalable Distributed Data Structures for Database Management

====
2001
====

2001-1 Silja Renooij (UU)
Qualitative Approaches to Quantifying Probabilistic Networks

2001-2 Koen Hindriks (UU)
Agent Programming Languages: Programming with Mental Models

2001-3 Maarten van Someren (UvA)
Learning as problem solving

2001-4 Evgueni Smirnov (UM)
Conjunctive and Disjunctive Version Spaces with Instance-Based Boundary Sets

2001-5 Jacco van Ossenbruggen (VU)
Processing Structured Hypermedia: A Matter of Style

2001-6 Martijn van Welie (VU)
Task-based User Interface Design

2001-7 Bastiaan Schonhage (VU)
Diva: Architectural Perspectives on Information Visualization

2001-8 Pascal van Eck (VU)
A Compositional Semantic Structure for Multi-Agent Systems Dynamics.

2001-9 Pieter Jan 't Hoen (RUL)
Towards Distributed Development of Large Object-Oriented Models, Views of Packages as Classes

2001-10 Maarten Sierhuis (UvA)
Modeling and Simulating Work Practice
BRAHMS: a multiagent modeling and simulation language for work practice analysis and design

2001-11 Tom M. van Engers (VUA)
Knowledge Management:
The Role of Mental Models in Business Systems Design

====
2002
====

2002-01 Nico Lassing (VU)
Architecture-Level Modifiability Analysis

2002-02 Roelof van Zwol (UT)
Modelling and searching web-based document collections

2002-03 Henk Ernst Blok (UT)
Database Optimization Aspects for Information Retrieval

2002-04 Juan Roberto Castelo Valdueza (UU)
The Discrete Acyclic Digraph Markov Model in Data Mining

2002-05 Radu Serban (VU)
The Private Cyberspace Modeling Electronic
Environments inhabited by Privacy-concerned Agents

2002-06 Laurens Mommers (UL)
Applied legal epistemology; Building a knowledge-based ontology of
the legal domain

2002-07 Peter Boncz (CWI)
Monet: A Next-Generation DBMS Kernel For Query-Intensive
Applications

2002-08 Jaap Gordijn (VU)
Value Based Requirements Engineering: Exploring Innovative
E-Commerce Ideas

2002-09 Willem-Jan van den Heuvel(KUB)
Integrating Modern Business Applications with Objectified Legacy
Systems

2002-10 Brian Sheppard (UM)
Towards Perfect Play of Scrabble

2002-11 Wouter C.A. Wijngaards (VU)
Agent Based Modelling of Dynamics: Biological and Organisational
Applications

2002-12 Albrecht Schmidt (Uva)
Processing XML in Database Systems

2002-13 Hongjing Wu (TUE)
A Reference Architecture for Adaptive Hypermedia Applications

2002-14 Wieke de Vries (UU)
Agent Interaction: Abstract Approaches to Modelling, Programming and Verifying Multi-Agent Systems

2002-15 Rik Eshuis (UT)
Semantics and Verification of UML Activity Diagrams for Workflow Modelling

2002-16 Pieter van Langen (VU)
The Anatomy of Design: Foundations, Models and Applications

2002-17 Stefan Manegold (UVA)
Understanding, Modeling, and Improving Main-Memory Database Performance

====
2003
====

2003-01 Heiner Stuckenschmidt (VU)
Ontology-Based Information Sharing in Weakly Structured Environments

2003-02 Jan Broersen (VU)
Modal Action Logics for Reasoning About Reactive Systems

2003-03 Martijn Schuemie (TUD)
Human-Computer Interaction and Presence in Virtual Reality Exposure Therapy

2003-04 Milan Petkovic (UT)
Content-Based Video Retrieval Supported by Database Technology

2003-05 Jos Lehmann (UVA)
Causation in Artificial Intelligence and Law - A modelling approach

2003-06 Boris van Schooten (UT)
Development and specification of virtual environments

2003-07 Machiel Jansen (UvA)
Formal Explorations of Knowledge Intensive Tasks

2003-08 Yongping Ran (UM)
Repair Based Scheduling

2003-09 Rens Kortmann (UM)
The resolution of visually guided behaviour

2003-10 Andreas Lincke (UvT)
Electronic Business Negotiation: Some experimental studies on the interaction between medium, innovation context and culture

2003-11 Simon Keizer (UT)
Reasoning under Uncertainty in Natural Language Dialogue using Bayesian Networks

2003-12 Roeland Ordelman (UT)
Dutch speech recognition in multimedia information retrieval

2003-13 Jeroen Donkers (UM)
Nosce Hostem - Searching with Opponent Models

2003-14 Stijn Hoppenbrouwers (KUN)
Freezing Language: Conceptualisation Processes across ICT-Supported Organisations

2003-15 Mathijs de Weerdt (TUD)
Plan Merging in Multi-Agent Systems

2003-16 Menzo Windhouwer (CWI)
Feature Grammar Systems - Incremental Maintenance of Indexes to Digital Media Warehouses

2003-17 David Jansen (UT)
Extensions of Statecharts with Probability, Time, and Stochastic Timing

2003-18 Levente Kocsis (UM)
Learning Search Decisions

====
2004
====

2004-01 Virginia Dignum (UU)
A Model for Organizational Interaction: Based on Agents, Founded in Logic

2004-02 Lai Xu (UvT)
Monitoring Multi-party Contracts for E-business

2004-03 Perry Groot (VU)
A Theoretical and Empirical Analysis of Approximation in Symbolic Problem Solving

2004-04 Chris van Aart (UVA)
Organizational Principles for Multi-Agent Architectures

2004-05 Viara Popova (EUR)
Knowledge discovery and monotonicity

2004-06 Bart-Jan Hommes (TUD)
The Evaluation of Business Process Modeling Techniques

2004-07 Elise Boltjes (UM)
Voorbeeldig onderwijs; voorbeeldgestuurd onderwijs, een opstap naar abstract denken, vooral voor meisjes

2004-08 Joop Verbeek(UM)
Politie en de Nieuwe Internationale Informatiemarkt, Grensregionale politi‰le gegevensuitwisseling en digitale expertise

2004-09 Martin Caminada (VU)
For the Sake of the Argument; explorations into argument-based reasoning

2004-10 Suzanne Kabel (UVA)
Knowledge-rich indexing of learning-objects

2004-11 Michel Klein (VU)
Change Management for Distributed Ontologies

2004-12 The Duy Bui (UT)
Creating emotions and facial expressions for embodied agents

2004-13 Wojciech Jamroga (UT)
Using Multiple Models of Reality: On Agents who Know how to Play

2004-14 Paul Harrenstein (UU)
Logic in Conflict. Logical Explorations in Strategic Equilibrium

2004-15 Arno Knobbe (UU)
Multi-Relational Data Mining

2004-16 Federico Divina (VU)
Hybrid Genetic Relational Search for Inductive Learning

2004-17 Mark Winands (UM)
Informed Search in Complex Games

2004-18 Vania Bessa Machado (UvA)
Supporting the Construction of Qualitative Knowledge Models

2004-19 Thijs Westerveld (UT)
Using generative probabilistic models for multimedia retrieval

2004-20 Madelon Evers (Nyenrode)
Learning from Design: facilitating multidisciplinary design teams

====
2005
====

2005-01Floor Verdenius (UVA)
Methodological Aspects of Designing Induction-Based Applications

2005-02 Erik van der Werf (UM))
AI techniques for the game of Go

2005-03 Franc Grootjen (RUN)
A Pragmatic Approach to the Conceptualisation of Language

2005-04 Nirvana Meratnia (UT)
Towards Database Support for Moving Object data

2005-05 Gabriel Infante-Lopez (UVA)
Two-Level Probabilistic Grammars for Natural Language Parsing

2005-06 Pieter Spronck (UM)
Adaptive Game AI

2005-07 Flavius Frasincar (TUE)
Hypermedia Presentation Generation for Semantic Web Information Systems

2005-08 Richard Vdovjak (TUE)
A Model-driven Approach for Building Distributed Ontology-based Web Applications

2005-09 Jeen Broekstra (VU)
Storage, Querying and Inferencing for Semantic Web Languages

2005-10 Anders Bouwer (UVA)
Explaining Behaviour: Using Qualitative Simulation in Interactive Learning Environments

2005-11 Elth Ogston (VU)
Agent Based Matchmaking and Clustering - A Decentralized Approach to Search

2005-12 Csaba Boer (EUR)
Distributed Simulation in Industry

2005-13 Fred Hamburg (UL)
Een Computermodel voor het Ondersteunen van Euthanasiebeslissingen

2005-14 Borys Omelayenko (VU)
Web-Service configuration on the Semantic Web; Exploring how semantics meets pragmatics

2005-15 Tibor Bosse (VU)
Analysis of the Dynamics of Cognitive Processes

2005-16 Joris Graaumans (UU)
Usability of XML Query Languages

2005-17 Boris Shishkov (TUD)
Software Specification Based on Re-usable Business Components

2005-18 Danielle Sent (UU)
Test-selection strategies for probabilistic networks

2005-19 Michel van Dartel (UM)
Situated Representation

2005-20 Cristina Coteanu (UL)
Cyber Consumer Law, State of the Art and Perspectives

2005-21 Wijnand Derks (UT)
Improving Concurrency and Recovery in Database Systems by Exploiting Application Semantics

====
2006
====

2006-01Samuil Angelov (TUE)
Foundations of B2B Electronic Contracting

2006-02 Cristina Chisăliță (VU)
Contextual issues in the design and use of information technology in organizations

2006-03 Noor Christoph (UVA)
The role of metacognitive skills in learning to solve problems

2006-04 Marta Sabou (VU)
Building Web Service Ontologies

2006-05 Cees Pierik (UU)
Validation Techniques for Object-Oriented Proof Outlines

2006-06 Ziv Baida (VU)
Software-aided Service Bundling - Intelligent Methods & Tools for Graphical Service Modeling

2006-07 Marko Smiljanic (UT)
XML schema matching -- balancing efficiency and effectiveness by means of clustering

2006-08 Eelco Herder (UT)
Forward, Back and Home Again - Analyzing User Behavior on the Web

2006-09 Mohamed Wahdan (UM)
Automatic Formulation of the Auditor's Opinion

2006-10 Ronny Siebes (VU)
Semantic Routing in Peer-to-Peer Systems

2006-11 Joeri van Ruth (UT)
Flattening Queries over Nested Data Types

2006-12 Bert Bongers (VU)
Interactivation - Towards an e-cology of people, our technological environment, and the arts

2006-13 Henk-Jan Lebbink (UU)
Dialogue and Decision Games for Information Exchanging Agents

2006-14 Johan Hoorn (VU)
Software Requirements: Update, Upgrade, Redesign - Towards a Theory of Requirements Change